Interior estimates for solutions of Abreu's equation

S. K. Donaldson

February 1, 2008

1 Introduction

In this paper we study a nonlinear, fourth order, partial differential equation for a convex function u on an open set Ω in $\mathbf{R}^n$. The equation can be written as

$$S(u) = A$$

where A is some given function and $S(u)$ denotes the expression

$$S(u) = -\sum_{i,j} \frac{\partial^2 u^{ij}}{\partial x^i \partial x^j}. \tag{1}$$

Here (u^{ij}) denotes the inverse of the Hessian matrix $u_{ij} = \frac{\partial^2 u}{\partial x^i \partial x^j}$. We call this PDE *Abreu's equation* since the expression $S(u)$ appears in [1], in the study of the differential geometry of toric varieties. In this paper we will be primarily interested in the case when A is a constant. Solutions to Equation 1 then correspond to certain Kahler metrics of constant scalar curvature, as we will recall in more detail in Section 5 below. Our purpose is to derive *a priori* estimates for solutions of Abreu's equation, which can be applied to existence questions for such constant scalar curvature metrics, on the lines of [6]. However in the present paper we will keep the differential geometry in the background, and concentrate on the PDE aspects of the equation.

Abreu's equation can be fitted, as a limiting case, into a class of equations considered by Trudinger and Wang [12],[13]. These authors study the Euler-Lagrange equation of the functional

$$\mathcal{J}_\alpha(u) = \int_\Omega (\det(u_{ij}))^\alpha,$$

(or this functional plus lower order terms). The Legendre transform interchanges the equations with parameters α and $1-\alpha$. There are two exceptional cases,

when $\alpha = 0$ or 1, when the Euler-Lagrange equations are trivial. Abreu's equation, which is the Euler-Lagrange equation associated to the functional

$$\int_\Omega -\log\det(u_{ij}) + Au,$$

is the natural limit of the Trudinger-Wang family when $\alpha \to 0$. Indeed, the Trudinger-Wang equations (in the absence of lower order terms) can be written as

$$\sum_{ij} \frac{\partial^2}{\partial x^i \partial x^j}(\det(u_{ij})^\alpha u^{ij}) = 0.$$

We will study Abreu's equation augmented by some specific boundary conditions. These depend on a measure σ on the boundary of Ω. We will consider two cases

Case 1 *Ω is the interior of a bounded polytope, defined by a finite number of linear inequalities with n codimension-1 faces of $\partial\Omega$ meeting at each vertex. The measure σ is a constant muptiple of the Lebesgue measure on each codimension-1 face of $\partial\Omega$.*

Case 2 *Ω is a bounded domain with strictly convex smooth boundary and σ is a smooth positive measure on $\partial\Omega$.*

The first case is the one which is relevant to toric varieties and is our main concern. We include the second case because it seems to lead to a natural PDE problem. In either case we define a class of convex functions $\mathcal{S}_{\Omega,\sigma}$ satisfying boundary conditions depending on σ. The detailed definitions are given in Section 2.2 below, but roughly speaking we require that a function u in $\mathcal{S}_{\Omega,\sigma}$ behaves like $\sigma^{-1} d\log d$ where d is the distance to the boundary and σ is regarded as a function, i.e. a multiple of the area measure on $\partial\Omega$. (Note here that the whole theory is affine-invariant, and does not depend on the choice of a Euclidean metric on $\mathbf{R}^n$, but for simplicity we will sometimes, as just above, express things in terms of the metric structure, although this is not playing any real role.) In this paper we study solutions u of Abreu's equation which lie in $\mathcal{S}_{\Omega,\sigma}$. As explained in [6] and in Section 2.2 below, these arise when one considers the problem of minimising the functional

$$\mathcal{F}(u) = -\int_\Omega \log\det u_{ij} + Au - \int_{\partial\Omega} u d\sigma, \tag{2}$$

over the set of smooth convex functions on Ω with L^1 boundary values. It is likely (although this requires proof) that any extrema of this functional lies in our space $\mathcal{S}_{\Omega,\sigma}$.

As explained in [6], simple examples show that for some data $\Omega, A, d\sigma$ there are no solutions of Abreu's equation with the given boundary behaviour. By

the same token, one needs further assumptions before *a priori* estimates of a solution can be obtained. The condition which we expect to be appropriate involves the linear part of the functional $\mathcal{F}$;

$$\mathcal{L}(f) = \int_{\partial\Omega} f d\sigma - \int_{\Omega} f A d\mu. \tag{3}$$

Fix a base point in the interior of Ω and call a function normalised if it vanishes, along with its first partial derivatives, at this base point. We consider the following condition on the data $(\Omega, A, d\sigma)$:

Condition 1 *The functional $\mathcal{L}$ vanishes on affine-linear functions and there is some $\lambda > 0$ such that*

$$\mathcal{L}(f) \geq \lambda^{-1} \int_{\partial\Omega} f d\sigma,$$

for all convex normalised functions f on $\overline{\Omega}$.

Of course when this condition holds we can fix $\lambda = \lambda(\Omega, \sigma, A)$ by taking the best possible constant. It is shown in [6], at least in the case when the dimension n is 2 and Ω is a polygon, that this is a necessary condition for the existence of a solution. Our goal in this paper is to derive *a priori* interior estimates for solutions assuming this condition. To streamline the statement we introduce the following terminology. We say a function $C(\Omega, \sigma, A, \lambda, d)$ where λ and d are positive real variables is *tame* if it is continuous with respect to the natural topology on the space of variables $(\Omega, \sigma, A, \lambda, d)$ (We use the $C^\infty_{\mathrm{loc}} \cap L^\infty$ topology on A; in Case 1 the space is divided into components labelled by the combinatorics of the faces and in Case 2 we use the C^∞ topology on Ω, σ). We use the same terminology for functions that depend on some subset of the variables. Our main result is

Theorem 1 *Suppose the dimension n is 2. There are tame functions K, C_p for $p = 0, 1, \ldots$ such that if A is a smooth bounded function in Ω, and (Ω, σ, A) satisfies Condition 1, then any normalised solution u in $\mathcal{S}_{\Omega,\sigma}$ of Abreu's equation satisfies*

$$K^{-1} \leq (u_{ij}) \leq K$$

and

$$|\nabla^p u| \leq C_p$$

where the argument d is the distance to the boundary of Ω and the argument λ is $\lambda(\Omega, \sigma, A)$

Of course we need not take the definition of "tame" functions $C(\Omega, \sigma, A, \lambda, d)$ too seriously. With a little labour we could make all of our estimates completely explicit. The definition is tailored to the continuity method: if we have a continuous 1-parameter family of such problems defined by data $(\Omega_t, \sigma_t, A_t)$

with solutions for $t < t_0$ then provided $\lambda(\Omega_t, \sigma_t, A_t)$ stays bounded the solutions cannot blow up in the interior as $t \to t_0$.

We present a variety of different arguments to establish these interior estimates. We hope this variety is justified by the desire to extend the results, in the future, in various directions: to the behaviour near the boundary and to higher dimensions. In Section 2 we bring together some more elementary preliminaries and in particular show that Condition 1 gives C^1 bounds on the solution. After this, the crucial intermediate goal is to obtain upper and lower bounds on the determinant of the Hessian (u_{ij}). In Section 3 we consider the case when the dimension n is 2 and A is a constant. We find a special argument in this case using a property of solutions of general elliptic equations in 2 dimensions. In Section 4 we find other arguments, using the maximum principle, which apply in any dimension. We get a lower bound on the determinant in all cases and, using the technique of Trudinger and Wang, an upper bound involving also a "modulus of convexity". As we explain in (5.1), in dimension 2, this modulus of convexity is controlled by the lower bound on the determinant, using an old result of Heinz. Once we have established upper and lower bounds on the determinant we can appeal to sophisticated analysis of Caffarelli and Gutiérrez to complete the proof of Theorem 1. This is explained in (5.1). In the case when A is a constant and Ω is a polygon in $\mathbf{R}^2$, which is our main interest, we give an alternative proof in the remainder of Section 5. This avoids the deep analysis of Caffarelli and Gutiérrez (and perhaps gives more explicit estimates) but uses their geometric results about the "sections" of a convex function in an essential way. The other ingredients are L^2 arguments, which make contact with Kahler geometry, a variant of Pogorelov's Lemma and standard linear theory (Moser iteration). We also explain that, in this case, one can avoid using Heinz's result, by combing results from Sections 3 and 4. In sum, in the case when A is constant and $\Omega \subset \mathbf{R}^2$ is a polygon we get a self-contained proof of Theorem 1, assuming only Chapter 3 of [9], and material from the textbook [7].

2 Preliminaries

2.1 Miscellaneous formulae

This subsection consists of entirely elementary material. We will give a number of useful equivalent forms of Abreu's equation, which are obtained by straightfoward manipulation. Throughout we use traditional tensor calculus notation, with summation convention.

Begin by considering any convex function u defined on an open set in $\mathbf{R}^n$. We write u_{ij} for the Hessian, u^{ij} for its inverse and U^{ij} for the matrix of cofactors i.e. $U^{ij} = \det(u_{ij})u^{ij}$. We can associate to u two second order, elliptic, linear differential operators

$$P(f) = \left(u^{ij} f_i\right)_j, \tag{4}$$

$$Q(f) = \left(U^{ij} f_i\right)_j . \tag{5}$$

Let $L = \log\det(u_{ij})$. Then we have identities

$$L_i = u^{ab} u_{abi}, \tag{6}$$

$$u_i^{jk} = -u^{ja} u_{abi} u^{bk}. \tag{7}$$

Thus

$$u^{ij} L_j = -u_j^{ij}. \tag{8}$$

This means, first, that

$$U_j^{ij} = \left(e^{-L} u^{ij}\right)_j = e^{-L} \left(u_j^{ij} - u^{ij} L_j\right) = 0.$$

Hence the operator Q can be written as

$$Q(f) = U^{ij} f_{ij}. \tag{9}$$

Now the first form of Abreu's equation, as in the Introduction is

$$u_{ij}^{ij} = -A.$$

We define a vector field $v = (v^j)$ by

$$v^j = -u_i^{ij} \tag{10}$$

So our second form of Abreu's equation is

$$v_j^j = A. \tag{11}$$

(Of course, the left hand side of this expression is the ordinary divergence of the vector field v.) On the other hand, by Equation 8, the vector field can also be expressed as

$$v^j = u^{ij} L_j,$$

so we have our third form of Abreu's equation

$$\left(u^{ij} L_i\right)_j = A, \tag{12}$$

that is

$$P(L) = A.$$

Expanding out the derivative we have

$$\left(u^{ij} L_i\right)_j = u^{ij} L_{,ij} + u_j^{ij} L_i = u^{ij} \left(L_{ij} - L_i L_j\right),$$

so we get our fourth form of the equation

$$u^{ij}(L_{ij} - L_i L_j) = A. \tag{13}$$

Finally, if we write

$$F = \det(u_{ij})^{-1} = e^{-L},$$

then

$$u^{ij} F_{ij} = u^{ij}(L_{ij} - L_i L_j) e^{-L} = -A e^{-L},$$

so we get our fifth form of the equation

$$Q(F) = -A. \tag{14}$$

2.2 The boundary conditions.

We begin by giving a precise definition of the set $\mathcal{S}_{\Omega,\sigma}$ of convex functions on Ω. The definitions are different in the two cases. We start with Case 1, when Ω is a polytope. For any point P of $\partial\Omega$ we can choose affine co-ordinates x_i on $\mathbf{R}^n$ such that P has co-ordinates $x_i = 0, i = 1, \dots n$ and a neighbourhood of p in $\overline{\Omega}$ is defined by p inequalities

$$x_1, x_2, \dots, x_p > 0.$$

We can also choose the co-ordinates so that the normal derivative of x_i on the face $x_i = 0$ of the boundary is σ^{-1}. We call such co-ordinates *adapted to* Ω *at* P.

Definition 1 *In Case 1 the set $\mathcal{S}_{\Omega,\sigma}$ consists of continuous convex functions u on $\overline{\Omega}$ such that*

- *u is smooth and strictly convex in Ω,*
- *The restriction of u to each face of $\partial\Omega$ is smooth and strictly convex;*
- *In a neighbourhood of any point P of $\partial\Omega$ the function u has the form*

  $$u = \sum x_i \log x_i + f$$

  where x_i are adapted co-ordinates, as above, and f is smooth up to the boundary.

(We say a smooth function is *strictly convex* if its Hessian is strictly positive.)

Now turn to Case 2, when Ω has smooth boundary. If P is a point of $\partial\Omega$ we can choose local co-ordinates $\xi, \eta_1, \dots \eta_{n-1}$ near P so that $\partial\Omega$ is given by the equation $\xi = 0$ and the normal derivative of ξ on the boundary is σ^{-1}. Again, we call such co-ordinates adapted. Define functions α_p of a positive real variable by

$$\alpha_1(t) = -\log t\ ,\ \alpha_2(t) = t^{-1}\ , \alpha_3(t) = t^{-2}.$$

Definition 2 *In Case 2 the set $\mathcal{S}_{\Omega,\sigma}$ consists of continous convex functions u on $\overline{\Omega}$ such that*

- *u is smooth and strictly convex on Ω;*
- *In a neighbourhood of any point P of $\partial\Omega$ there are adapted co-ordinates $(\xi, \underline{\eta_i})$ in which*
$$u = \xi \log \xi + f$$
where for $p \geq 1$, $p + q \leq 3$
$$|\nabla_\xi^p \nabla_{\underline{\eta}}^q f| = o(\alpha_p(\xi)).$$
as $\xi \to 0$.

Here the notation $\nabla_{xi}^p \nabla_{\underline{\eta}}^q$ means any partial derivative of order p in the variable ξ and total order q in the η_i.

Now, for small positive δ, let $\Omega_\delta \subset \Omega$ be the set of points distance at least δ from $\partial\Omega$, i.e. a "parallel" copy of $\partial\Omega$. If $u \in \mathcal{S}_{\Omega,\sigma}$ and χ is a smooth function on Ω, integration-by-parts over Ω_δ gives the fundamental identity:

$$\int_{\Omega_\delta} u^{ij}\chi_{ij} = \int_{\Omega_\delta} u^{ij}_{ij}\chi + \int_{\partial\Omega_\delta} -u^{ij}_j \chi + u^{ij}\chi_j. \tag{15}$$

We can write the boundary terms as

$$\int_{\partial\Omega_\delta} \chi \, v_{\text{norm}} + \nabla_X \chi,$$

where v is the vector field introduced in (10) above, v_{norm} is its normal component and X is the vector field, defined on a neighborhood of $\partial\Omega$ by

$$X^j = u^{ij}\nu_j,$$

ν_j being the unit normal to $\partial\Omega$ at the closest boundary point. (In Case 1, the vector field X will be discontinuous near the "corners" of $\partial\Omega$ but this will not matter.)

The main result we need is

Proposition 1 *In either case, if $u \in \mathcal{S}_{\Omega,\sigma}$ then as $\delta \to 0$; $|X| = O(\delta)$ and v_{norm} converges uniformly to σ.*

Here $|X|$ refers to the Euclidean length and we interpret v_{norm} as a function on $\partial\Omega$ in the obvious way, by taking the closest point of $\partial\Omega_\delta$. We assume Proposition 1 for the moment. Taking the limit as δ tends to 0 in Equation 15, we obtain

Corollary 1 *Let u be in $\mathcal{S}_{\Omega,\sigma}$ with $S(u) = -u^{ij}_{ij} \in L^\infty(\Omega)$ and let χ be a continuous, convex function on $\overline{\Omega}$ smooth in the interior and with $\nabla\chi = o(d^{-1})$, where d is the distance to $\partial\Omega$. Then $u^{ij}\chi_{ij}$ is integrable in Ω and*

$$\int_\Omega u^{ij}\chi_{ij} = \int_\Omega u^{ij}_{ij}\chi + \int_{\partial\Omega} \chi d\sigma.$$

The main application we make of Corollary 1 is the case when $\chi = u$. It is clear from the definitions that ∇u is $O(-\log d)$, hence $o(d^{-1})$, near the boundary, and $u_{ij}u^{ij} = n$. So we have the identity

$$\mathcal{L}u = n$$

where $A = -u^{ij}_{ij}$ and $\mathcal{L}$ is the linear functional defined in (3). Thus we obtain

Corollary 2 *Suppose $A \in L^\infty(\Omega)$ and that $\mathcal{L}$ satisfies Condition 1. Then if u is a normalised function in $\mathcal{S}_{\Omega,\sigma}$ which satisfies Abreu's equation $S(u) = A$ we have*

$$\int_{\partial\Omega} u \; d\sigma \leq n\lambda.$$

A simple argument ([6], Lemma 5.2.3) shows that the integral over the boundary, for a normalised convex function, controls the derivative in the interior and we have

Corollary 3 *Under the same hypotheses as Corollary 2,*

$$|\nabla u| \leq C d^{-n},$$

where C depends tamely on $(\Omega, \sigma, \lambda)$.

This first derivative bound is the seed which we wish to develop in this paper to obtain bounds on higher derivatives.

We mention some other applications of Corollary 1. Here we will restrict for simplicity to the case when Ω is a polytope.

Corollary 4 *In the case when $\Omega \subset \mathbf{R}^n$ is a polytope, for any $u \in \mathcal{S}_{\Omega,\sigma}$, $|\log\det u_{ij}|$ is integrable over Ω so for any $A \in L^\infty(\Omega)$ the functional $\mathcal{F}_A$ is defined on $\mathcal{S}_{\Omega,\sigma}$. The functional $\mathcal{F}_A$ is convex on $\mathcal{S}_{\Omega,\sigma}$ and the equation $S(u) = A$ has at most one normalised solution u in $\mathcal{S}_{\Omega,\sigma}$. For such u,*

$$\int_\Omega |\log\det(u_{ij})| \leq C,$$

where C is a tame function of $\Omega, \sigma, \|A\|_{L^\infty}, \lambda$.

The proof of this Corollary follows easily from the results in ([6], subsections (3.3) and (5.1)). The restriction to Case 1 arises because in this case, as is will be clear from Proposition 5 below, $S(u)$ is bounded for any $u \in \mathcal{S}_{\Omega,\sigma}$. In Case 2 we have not analysed the behaviour of $S(u)$ near the boundary, for general elements of $\mathcal{S}_{\Omega,\sigma}$, and we leave this issue to be discussed elsewhere.

2.3 Proof of Proposition 1.

In the course of the proof we will also establish some other properties of functions in $\mathcal{S}_{\Omega,\sigma}$. Before beginning it is worth pointing that, in a sense the proofs of these boundary properties need not be taken too seriously at this stage of the development of the theory. At this stage we are in essence free to choose the definition of $\mathcal{S}_{\Omega,\sigma}$ and we could impose any reasonable conditions we choose. For example we we could take the conclusions of Proposition 1 as part of the *definition* of $\mathcal{S}_{\Omega,\sigma}$. The discussion will only acquire an edge when one goes on to the *existence* theory. It is possible that the detailed definitions may need to be modified then. At the present stage, the definitions we have concocted serve to indicate at least the nature of the solutions we want to consider while being, we hope, sufficiently general to permit a sensible existence theory.

We begin with Case 1. This is not very different from the analysis of metrics on toric varities in [1], [8], but we include a discussion for completeness. We work in adapted co-ordinates around a boundary point, so

$$u = \sum_{i=1}^{p} x_i \log x_i + f$$

where f is smooth up to the boundary. By the second condition of Definition 1 we may choose the co-ordinates so that at $\underline{x} = 0$

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \delta_{ij},$$

for $p+1 \leq i, j \leq n$. Thus

$$(u_{ij}) = \mathrm{diag}(x_1^{-1}, \dots, x_p^{-1}, 1 \dots, 1) + (\psi_{ij}), \tag{16}$$

where ψ_{ij} are smooth up to the boundary and ψ vanishes at $\underline{x} = 0$ for $i, j \geq p+1$.

Proposition 2

- $\det(u_{ij}) = x_1 \dots x_p \Delta$ *where* Δ *is smooth up to the boundary and* $\Delta(0) = 1$.
- *There are functions* f_i, g_{ij}, h_{ij}, *all smooth up to the boundary and with* $f_i(0) = 1$, $h_{ij}(0) = 0$, *such that*

$$(u_{ij}) = \mathrm{diag}(f_1 x_1, \dots, f_p x_p, f_{p+1}, \dots f_n) + \left(\sigma^{ij}\right),$$

where

$$\sigma^{ij} = x_i x_j g_{ij}$$

for $1 \leq i, j \leq p$;

$$\sigma^{ij} = x_i g_{ij}$$

for $1 \leq i \leq p$, $j > p$;

$$\sigma^{ij} = h_{ij}$$

for $i, j > p$.

It is completely straightforward to verify Proposition 1, in Case 1, given this Proposition 2. Notice that, according to Proposition 2, the matrix u^{ij} is smooth up to the boundary.

Let Λ be the diagonal matrix

$$\Lambda = \mathrm{diag}(x_1^{1/2}, \dots x_p^{1/2}, 1 \dots, 1).$$

To prove the first item of Proposition 2 we write H for the Hessian matrix (u_{ij}) and consider the matrix $\Lambda^2 H$. This has the form $1+E$ where the entries E_{ij} of the matrix E are smooth up to the boundary and E_{ij} vanishes at 0 except for the range $1 \leq j \leq p,\ i > p$. Thus, at 0, the matrix E is strictly lower-triangular and so $\Delta = \det(1+E)$ is a smooth function taking the value 1 at $\underline{x} = 0$, and

$$\det(u_{ij}) = \det \Lambda^{-2}\ \Delta = x_1^{-1} \dots x_p^{-1} \Delta.$$

To prove the second item we consider the symmetric matrix $\Lambda H \Lambda$. We write $\Lambda H \Lambda = 1 + (F_{ij})$ and Equation 16 yields

$$F_{ij} = A_{ij} x_i^{1/2} x_j^{1/2}$$

for $1 \leq i, j \leq p$;

$$F_{ij} = F_{ji} = B_{ij} x_i^{1/2},$$

for $1 \leq i \leq p$ and $j > p$ where A_{ij}, B_{ij} are smooth up to the boundary. In the remaining block, $p+1 \leq i, j \leq n$, the F_{ij} are smooth up to the boundary, vanishing at 0. Let $\mathrm{ad}(\Lambda H \Lambda)$ be the matrix of co-factors, or "adjugate matrix", so that

$$(\Lambda H \Lambda)^{-1} = \det(\Lambda H \Lambda)^{-1} \mathrm{ad}(\Lambda H \Lambda).$$

We claim that $\mathrm{ad}(\Lambda H \Lambda)$ has the form $1 + F'_{ij}$ where F' satisfies the same conditions as F above; that is,

$$F'_{ij} = A'_{ij} x_i^{1/2} x_j^{1/2}$$

for $1 \leq i, j \leq p$;

$$F'_{ij} = B'_{ij} x_i^{1/2}$$

for $1 \leq i \leq p$ and $j > p$ etc. The proof of this is a straightforward matter of considering the various terms in the cofactor determinants which we leave to the reader. Given this, the second item of the Proposition follows by writing

$$\left(u^{ij}\right) = H^{-1} = \Delta^{-1}\ \Lambda\ \mathrm{ad}(\Lambda H \Lambda)\ \Lambda.$$

We now turn to Case 2. For simplicity we will consider the case when $n = 2$ (see the remarks at the beginning of this subsection). Moreover, to make the calculations easier, we will consider a special kind of adapted co-ordinate. Suppose the point P is the origin and let the boundary of Ω be represented by

a graph $x_1 = Q(x_2)$ where $Q(0) = Q'(0) = 0, Q''(0) < 0$. Then we take the adapted co-ordinates

$$\eta = x_2 \ , \ \xi = \rho(\eta)(Q(x_2) - x_1),$$

for a positive function ρ equal to σ^{-1}, where σ is regarded as a function of x_2 using the obvious parametrisation of the boundary. Without loss of generality we can compute at points on the x_2 axis, where $\eta = 0$. We write $\partial_{x_1}, \partial_{x_2}$ for $\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}$ and $\partial_\xi, \partial_\eta$ for $\frac{\partial}{\partial \xi}, \frac{\partial}{\partial \eta}$. Then, transforming between the co-ordinate systems (x_1, x_2) and (ξ, η), we have

$$\partial_{x_1} = \rho \partial_\xi \ , \ \partial_{x_2} = \partial_\eta + F\partial_\xi,$$

where $F(\xi, \eta)$ has the form $F = \xi A(\eta) + B(\eta)$ and $B(0) = 0, B'(0) < 0$. Thus

$$\begin{aligned} \partial_{x_1}^2 &= \rho^2 \partial_\xi^2, & (17)\\ \partial_{x_1}\partial_{x_2} &= \rho^2 \partial_\xi \partial_\eta + \rho F \partial_\xi^2 + \rho A \partial_\xi, & (18)\\ \partial_{x_2}^2 &= (\partial_\eta^2 + 2F\partial_\xi\partial_\eta + F^2\partial_\xi^2) + (\xi A' + B' + A(\xi A + B))\partial_\xi. & (19) \end{aligned}$$

Applying this to a function $u = (\xi \log \xi - \xi) + f$ in $\mathcal{S}_{\Omega,\sigma}$ (where we have replaced f by $f - \xi$ in Definition 2, which obviously makes no difference) we get

$$(u_{ij}) = \begin{pmatrix} \rho^2\xi^{-1} & 0 \\ 0 & B'\log\xi \end{pmatrix} + \begin{pmatrix} \alpha & \beta \\ \beta & \gamma \end{pmatrix},$$

where

$$\alpha = \partial_{x_1}^2 f \ , \ \beta = \rho A \log \xi + \partial_{x_1}\partial_{x_2} f \ , \ \gamma = (\xi A' + B' + A(\xi A + B))\log\xi + \partial_{x_2}^2 f.$$

Then

$$\det(u_{ij}) = (\rho^2 B')\xi^{-1}\log\xi \ \Delta,$$

where

$$\Delta = \left(1 + \frac{\alpha\xi}{\rho^2}\right)\left(1 + \frac{\gamma}{B'\log\xi}\right) - \frac{\beta^2\xi}{\rho^2 B'\log\xi}.$$

Write

$$\lambda = \frac{\alpha\xi}{\rho^2}, \mu = \frac{\gamma}{B'\log\xi}, \nu = \frac{\beta\xi}{B'\rho^2\log\xi}.$$

Then

$$\Delta = (1+\lambda)(1+\mu) - \beta\nu,$$

and the inverse matrix is

$$\left(u^{ij}\right) = \Delta^{-1}\begin{pmatrix} \frac{\xi}{\rho^2}(1+\mu) & -\nu \\ -\nu & \frac{1}{B'\log\xi}(1+\lambda) \end{pmatrix}.$$

Now first we evaluate at a point where $\eta = 0$, so $B = 0$. Then we have

$$\alpha = \rho^2 \partial_\xi \partial_\xi f,$$

and by the definition of $\mathcal{S}_{\Omega,\sigma}$ this is $o(\xi^{-1})$. So λ is $o(1)$. We have

$$\beta = \rho^2 \partial_\xi \partial_\eta f + \rho \xi A \partial_\xi \partial_\xi f + \rho A (\partial_\xi f + \log \xi),$$

and applying the definition of $\mathcal{S}_{\Omega,\sigma}$ to the various terms we see that $\beta = O(|\log \xi|)$. This means that ν is $O(\xi)$ and $\beta\nu$ is $O(\xi|\log \xi|)$, which are both $o(1)$. Similarly, applying the definitions, we find that μ is $o(1)$. We see from this first that the vector field Z, which has components u^{11}, u^{12} is $O(\xi)$, so verifying the first item of Proposition 1. We also see that

$$\det(u_{ij}) \sim \frac{\rho^2}{B'} \xi^{-1} |\log \xi|, \tag{20}$$

since $\Delta \sim 1$. To complete the proof we need to differentiate again. We have

$$L = \log \det(u_{ij}) = -\log \xi + \log(-\log \xi) + \log(-B') + \log \rho^2 + \log \Delta,$$

Thus, when $\eta = 0$,

$$\begin{aligned} \partial_{x_1} L &= \rho(\xi^{-1} + (\xi \log \xi)^{-1}) + \rho(\frac{\partial_\xi \Delta}{\Delta}), & (21) \\ \partial_{x_2} L &= \xi A \partial_\xi L + \frac{\partial_\eta \Delta}{\Delta}. & (22) \end{aligned}$$

We claim that $\partial_\eta \Delta$ is $O(1)$ and $\partial_\xi \Delta$ is $o(\xi^{-1})$. Given this, we have

$$v^1 = \Delta^{-1}(\rho^{-1}(1+\mu) + o(\xi)), v^2 = \Delta^{-1}(-\nu\rho\xi^{-1} + O(|\log \xi|^{-1}),$$

and we see that the normal component v^1 converges to $\sigma = \rho^{-1}$ as desired. To verify the claim we have to show that $\partial_\eta \lambda, \partial_\eta \mu, \partial_\eta(\beta\nu)$ are all $O(1)$ and $\partial_\xi \lambda, \partial_\xi \mu, \partial_\xi(\beta\nu)$ are all $o(\xi^{-1})$. This is just a matter of differentiating the formulae defining λ, μ, ν and applying the definition of $\mathcal{S}_{\Omega,\sigma}$ to each term: we omit the details.

3 The two dimensional case.

3.1 The conjugate function

Throughout this Section 3 we suppose that A is a constant and u is a solution of Abreu's equation in Ω. We begin with second form of the equation

$$div(v) = v^i_i = A.$$

The "radial" vector field x^i on $\mathbf{R}^n$ has divergence n, so if we define a vector field w by

$$w^i = v^i - \frac{A}{n} x^i, \tag{23}$$

then $w^i_i = 0$: the vector field w has divergence zero. Now define a function h by

$$h = u - u_k x^k. \tag{24}$$

Then

$$h_i = u_i - u_{ki} x^k - u_k \delta^k_i = -u_{ki} x^k, \tag{25}$$

so

$$u^{ij} h_i = -x^j. \tag{26}$$

Thus, using Equation 8, we have

$$w^i = u^{ij} \tilde{L}_j, \tag{27}$$

where $\tilde{L} = L + \frac{A}{n} h$.

We now specialise (for the rest of this subsection) to the case when $n = 2$. The special feature here is that divergence-free vector fields can be represented by Hamiltonians, so there is a function H (unique up to a constant) with

$$w^i = \epsilon^{ij} H_j,$$

where ϵ^{ij} is the skew-symmetric tensor with $\epsilon^{12} = 1$.

Lemma 1 *The function H satisfies the equation $Q(H) = 0$.*

To see this we write

$$\tilde{L}_b = -u_{bk} \epsilon^{kj} H_j.$$

Then we have

$$\tilde{L}_{ab} \epsilon^{ab} = 0$$

by the symmetry of second derivatives, whereas, differentiating Equation 27,

$$\tilde{L}_{ab} \epsilon^{ab} = -\left(\epsilon^{ab} u_{bk} \epsilon^{kj} H_j \right)_a = Q(H),$$

since $U^{aj} = \epsilon^{ab} u_{bk} \epsilon^{kj}$.

We call H the *conjugate function* since the relationship between the functions H and $\tilde{L}$, is analogous to that between conjugate harmonic functions in two dimensions, except that they satisfy *different* linear elliptic equations: $P(\tilde{L}) = 0, Q(H) = 0$.

Next we use the boundary conditions, so we suppose that u is in $\mathcal{S}_{\Omega,\sigma}$. We have shown in Proposition 1 that the normal component of the vector field v on the boundary can be identified with given measure σ. Likewise, the vector

field $\frac{A}{2}x^i$ defines a (signed) measure $d\tau$ on the boundary. The condition that $\mathcal{L}$ vanishes on the constants implies that

$$\mathrm{Vol}(\partial\Omega, \sigma) = A\mathrm{Vol}(\Omega), \tag{28}$$

which in turn implies

$$\int_{\partial\Omega} d\sigma - d\tau = 0.$$

Thus there is a function b on $\partial\Omega$, unique up to a constant, with $db = \sigma - \tau$. Observe that, in the case when Ω is a polygon, the function b is linear on each face of $\partial\Omega$. For in this case $d\sigma$ is constant multiple of the Lebesgue measure on each face and the *normal* component of the radial field is also constant on each face.

Lemma 2 *The function H extends continously to $\overline{\Omega}$ with boundary value b (up to the addition of a constant).*

First suppose that H extends smoothly up to the boundary. The relation $w^j = \epsilon^{ij} H_i$ asserts that the normal component of w is equal to the tangential derivative of H. But the normal component of w is the derivative of b, so the derivative of $H - b$ vanishes. For the general case we apply the same argument to a slightly smaller domain and take a limit, the details are straightforward.

We can now apply a special result ([7], Lemma 12.6) for solutions of elliptic equations

$$\sum_{i,j=1}^{2} a^{ij} \frac{\partial^2 f}{\partial x^i \partial x^j} = 0$$

in two dimensions whose boundary values satisfy a "three point" condition. Here we use Equation 9 to see that the equation $Q(H) = 0$ has this form, with $a^{ij} = U^{ij}$. The three point condition requires that for any three points X_1, X_2, X_3 on $\partial\Omega$ the slope of the plane in $\mathbf{R}^3$ containing $(X_i, H(X_i))$ is bounded by some fixed K. This holds in our situation because the restriction of H is b, which is smooth in the strictly convex case and linear-on-faces in the polygonal case. In fact, a little reflection shows that in Case 1 we can take K to be the maximum of the slopes attained by taking X_1, X_2, X_3 to be vertices of the polygon. So we can apply the result of [7] quoted above, and deduce that the derivative of H is bounded by the constant throughout Ω. Thus the vector field w satisfies an L^∞ bound over Ω and hence the same is true of v. To sum up we have

Theorem 2 *Suppose $u \in \mathcal{S}_{\Omega,\sigma}$ is a solution of Abreu's equation where A is a constant and the dimension n is 2. Then there is a constant K, depending tamely on Ω, σ, such that $|v| \leq K$ throughout Ω.*

3.2 Integrating bounds on v

In this subsection we study the implications for $L = \log\det u_{ij}$ of an L^∞ bound on the vector field $v^i = -u^{ij}_j$ associated to a convex function u on a convex domain $D \subset \mathbf{R}^n$. We prove two results, under different hypotheses. For the first, recall that the derivative ∇u is a diffeomorphism from D to its image $(\nabla u)(D) \subset \mathbf{R}^n$.

Lemma 3 *Suppose that the image $(\nabla u)(D)$ is convex and $|v| \leq K$ on D. Then for any $x, y \in D$*

$$|L(x) - L(y)| \leq K|(\nabla u)(x) - (\nabla u)(y)|.$$

To see this, let ξ_i be the standard Euclidean co-ordinates on the image $(\nabla u)(D)$. (More invariantly, ∇u should be thought of as mapping D to the dual space $(\mathbf{R}^n)^*$ and ξ_i are just the dual co-ordinates of x^i.) In traditional notation, $\xi_i = u_i$ so

$$\frac{\partial \xi_i}{\partial x^j} = u_{ij} \; , \frac{\partial x^j}{\partial \xi_i} = u^{ij}.$$

Now $v^i = u^{ij} L_j$ (see (8)) but by the chain rule

$$u^{ij} L_j = \frac{\partial L}{\partial \xi_i}.$$

In other words, if we define a function L^* on $(\nabla u)(D)$ to be the composite $L \circ (\nabla u)^{-1}$ then v is the derivative of L^* and hence $|\nabla L^*| \leq K$. If $(\nabla u)(D)$ is convex then $\nabla u(x), \nabla u(y)$ can be joined by a line segment in $\nabla u(D)$ and the result follows immediately by integrating the derivative bound along the segment.

The next results applies to solutions of Abreu's equation, but avoids the convexity hypothesis on the image under ∇u. We write $\mathrm{Av}_D(L)$ for the mean value

$$\mathrm{Av}_D(L) = (\mathrm{Vol}(D))^{-1} \int_D L.$$

Theorem 3 *Suppose that $D \subset \mathbf{R}^n$ is a bounded convex domain with smooth boundary and $x_0 \in D$ is a base point. Let R be the maximal distance from x_0 to a point of ∂D. Let u be a solution of Abreu's equation $S(u) = A$ in D, smooth up to the boundary and normalised at x_0. Then*

$$|L(x_0) - \mathrm{Av}_D(L)| \leq C \int_{\partial D} u \, d\nu,$$

where $d\nu$ is the standard Riemannian volume form on ∂D and

$$C = (n\mathrm{Vol}(D))^{-1} \left(n \sup_D |v| + R \sup_D |A| \right).$$

To prove this, we can suppose that x_0 is the origin and let $(r, \underline{\theta})$ denote standard "polar-coordinates" on $\mathbf{R}^n \setminus \{0\}$ (so $\underline{\theta} \in S^{n-1}$). Suppose that ∂D is given in these co-ordinates by an equation $r = \rho(\underline{\theta})$. Let f be the function

$$f = \left(1 - \frac{\rho(\underline{\theta})^n}{r^n}\right)$$

on $\mathbf{R}^n \setminus \{0\}$ and define a vector field ζ on $\mathbf{R}^n \setminus \{0\}$ by

$$\zeta^i = f x^i.$$

Then, away from the origin,

$$\zeta^i_i = x^i_i(1 - \frac{\rho^n}{r^n}) - x^i \frac{\partial}{\partial x^i}\frac{\rho^n}{r^n} = n(1 - \frac{\rho^n}{r^n}) - r\frac{\partial}{\partial r}\frac{\rho^n}{r^n} = n,$$

and ζ vanishes on the boundary of D. It is clear that ζ satisfies the distributional equation

$$\zeta^i_i = n - n\mathrm{Vol}(D)\ \delta_0,$$

where δ_0 is the delta-function at the origin. Applying this distribution to the function L we get

$$n(\mathrm{Vol}(D)L(0) - \int_D L) = \int_D \zeta^i L_i.$$

Now recall that $h = u - u_i x^i$ satisfies $h_i = -u_{ij}x^j$. Then $L_i = u_{ij}v^j$, and

$$x^i L_i = u_{ij}x^i v^j = -h_j v^j = -\left(hv^j\right)_j + hv^j_j.$$

Using Abreu's equation in the form $v^j_j = A$ we have

$$x^i L_i = -\left(hv^j\right)_j + Ah.$$

Now

$$\int_D \zeta^i L_i = \int_D f x^i L_i = -\int_D f(hv^j)_j + \int_D Afh.$$

We can integrate by parts on the first term to get

$$\int_D \zeta^i L_i = \int_D (hf_j v^j + Afh).$$

There is no boundary term from $\partial\Omega$ since f vanishes on the boundary. There is also no extra term caused by the singularity of f at the origin since h vanishes to second order there. To sum up, we have the identity,

$$n\mathrm{Vol}(D)\,(L(0) - \mathrm{Av}_D(L)) = (I) + (II), \tag{29}$$

where

$$(I) = \int_D h f_j v^j \qquad (II) = \int_D A f h.$$

We next estimate the size of the integrals (I) and (II). We have

$$\frac{\partial f}{\partial r} = -n\frac{\rho^n}{r^{n+1}} \quad , \quad \frac{1}{r}\frac{\partial f}{\partial \underline{\theta}} = n\frac{\rho^{n-1}}{r^{n+1}}\frac{\partial \rho}{\partial \underline{\theta}},$$

so

$$|\nabla f| = \frac{n\rho^{n-1}}{r^{n+1}}W(\underline{\theta})$$

where

$$W(\theta) = \sqrt{\rho^2 + |\rho^2_{\underline{\theta}}}.$$

Now let $K = \sup_D |v|$ so we have

$$|f_j v^j| \leq \frac{nK\rho^{n-1}}{r^{n+1}}W(\underline{\theta}).$$

Thus

$$|(I)| \leq nK\int_D W(\underline{\theta})\rho(\underline{\theta})^{n-1}\frac{h}{r^{n+1}} = 2K\int_D \rho^{n-1}W(\underline{\theta})\frac{h}{r^2}drd\underline{\theta},$$

in an obvious notation. Now consider this integral along a fixed ray $\underline{\theta}$ = constant. In polar co-ordinates we can write

$$h = u - r\frac{\partial u}{\partial r},$$

so

$$\frac{\partial}{\partial r}(r^{-1}u) = -r^{-2}u + r^{-1}\frac{\partial u}{\partial r} = -r^{-2}h.$$

So

$$\int_\epsilon^{\rho(\underline{\theta})}\frac{h(r)}{r^2}dr = \rho^{-1}u(\rho(\underline{\theta}),\underline{\theta}) - \epsilon^{-1}u(\epsilon,\underline{\theta}),$$

and the term from the lower limit tends to zero with ϵ since the derivative of u vanishes at the origin. So we have

$$|(I)| \leq nK\int_{\partial D}\rho(\underline{\theta})^{n-2}W(\theta)ud\underline{\theta}. \tag{30}$$

For the other integral we have $-ru_r \leq h \leq 0$, and $f \leq 0$ so

$$0 \leq (II) \leq \sup|A|\int_D \frac{\rho^n}{r^n}ru_r r^{n-1}drd\underline{\theta} = \sup|A|\int_{\partial D}\rho(\underline{\theta})^n ud\underline{\theta}. \tag{31}$$

We now transform the integrals over ∂D to the Riemannian area form $d\nu$. By straightforward calculus

$$d\nu = \rho^{n-1}\sqrt{1+\left(\frac{\rho_{\underline{\theta}}}{\rho}\right)^2}\, d\underline{\theta} = \rho^{n-2}W(\underline{\theta})d\underline{\theta}.$$

So inequality 30 is just

$$|(I)| \leq nK \int_{\partial D} u d\nu.$$

Also, $d\nu \geq \rho^{n-1} d\theta$ so inequality 31 gives

$$|(II)| \leq \sup|A| \int_{\partial D} u\rho d\nu \leq R \sup|A| \int_{\partial D} u d\nu,$$

and putting these together gives the result stated.

We now apply these results in the case when $n = 2$ and A is a constant, using Theorem 2. We obtain

Theorem 4 *Let $\Omega \subset \mathbf{R}^2$ and $u \in \mathcal{S}_{\Omega,\sigma}$ be a normalised solution of Abreu's equation with A constant. Suppose Condition 1 holds. Then*

$$L \leq C_0 + C_1|\nabla u| \leq C_2 d^{-2},$$

where d is the distance to the boundary of Ω and C_0, C_1, C_2 depends tamely on Ω, σ, λ.

We can prove this using the simpler result Lemma 3. Any function u in $\mathcal{S}_{\Omega,\sigma}$ is continuous on $\overline{\Omega}$ and so attains a minimum value on $\overline{\Omega}$, but it is clear from the definition of $\mathcal{S}_{\Omega,\sigma}$ that this cannot be attained on the boundary. Changing u by the addition of any linear function we see that for $u \in \mathcal{S}_{\Omega,\sigma}$ the image $(\nabla u)(\Omega)$ is the whole of $\mathbf{R}^n$. Let B be a small disc about the base point. We know that ∇u is bounded on B by Corollary 3, and this gives a bound on the area of $\nabla u(B)$, which is

$$\int_B \det u_{ij}.$$

So there is a c, depending tamely on the data, such that we can find some point in B where $L \leq c$. Now the result follows by combining Corollary 3, Theorem 2 and Lemma 3.

We can give a slightly different proof of the above result using Theorem 3 in place of Lemma 3. In the case when $\Omega \subset \mathbf{R}^2$ is a polygon and A is a constant we can also obtain an interior lower bound on L by applying Corollary 4, which gives a bound on

$$\int_\Omega |L|.$$

However we will not discuss this in detail, since we will get a better lower bound in the next section.

4 Applications of the maximum principle

4.1 Lower bound for the determinant

In this Section we will derive a lower bound for $L = \log\det(u_{ij})$, valid in any dimension and for any bounded function A.

Theorem 5 *Suppose that $u \in \mathcal{S}_{\Omega,\sigma}$ is a solution of Abreu's equation for some bounded function A. For any $\alpha \in (0,1)$ there is a constant C_α (depending only on n and α) such that*

$$\det u_{ij} \geq C_\alpha d^{-\alpha} (\sup A)^n \ \mathrm{Diam}(\Omega)^{2n+\alpha}$$

throughout Ω.

Here we recall that d is the distance to the boundary of Ω. Diam (Ω) is the diameter of Ω.

We can compare this result with the *asymptotic behaviour* established in Section 2 (Proposition 2 and (20)). In Case 1

$$\det u_{ij} \sim C_1 d^{-1}, \tag{32}$$

as x tends to a generic boundary point (i.e on an $(n-1)$-dimensional face of the boundary) and in Case 2,

$$\det u_{ij} \sim C_2 d^{-1} |\log d|. \tag{33}$$

To prove the theorem we suppose that we have a function ψ on Ω which satisfies the following conditions.

1. For each point of Ω the matrix

$$M_{ij} = \psi_{ij} - \psi_i \psi_j$$

   is positive definite.

2. $L - \psi$ tends to $+\infty$ on $\partial\Omega$.

Certainly such functions exist. For example, we can take $\psi_\epsilon(x) = \epsilon |x|^2$ for small ϵ: the second condition holds since $L \to \infty$ on $\partial\Omega$. Given such a function ψ, we write $D = \det(M_{ij})$.

By the second condition there is a point p in Ω where $L - \psi$ attains its minimum value. At this point we have

$$L_i = \psi_i \quad (L_{ij} - \psi_{ij}) \geq 0.$$

We take Abreu's equation in the form (Equation 13)

$$u^{ij}(L_{ij} - L_i L_j) = A.$$

At the point p we have

$$u^{ij}L_{ij} = A + u^{ij}L_iL_j = A + u^{ij}\psi_i\psi_j,$$

and

$$u^{ij}L_{ij} \geq u^{ij}\psi_{ij}.$$

Thus, at the point p,

$$u^{ij}(\psi_{ij} - \psi_i\psi_j) = u^{ij}M_{ij} \leq A.$$

So $u^{ij}M_{ij} \leq \overline{A}$, where $\overline{A} = \sup_\Omega A$. We now use the standard inequality for positive definite matrices,

$$\left(\det(u^{ij})\det(M_{ij})\right)^{1/n} \leq n^{-1}u^{ij}M_{ij}. \tag{34}$$

(This is just the arithmetic-geometric mean inequality for the relative eigenvalues of $(u_{ij}), (M_{ij})$.) Now use the fact that $\det(u^{ij}) = e^{-L}$ to get, at the point p,

$$e^{-L}D \leq \left(\overline{A}/n\right)^n,$$

or

$$L(p) \geq \log D(p) - c,$$

where $c = n\log(\overline{A}/n)$. Then for any other point, q, in Ω we have $(L-\psi)(q) \geq (L-\psi)(p)$ so

$$L(q) \geq L(p) + \psi(q) - \psi(p) \geq \log D(p) + \psi(q) - \psi(p) - c,$$

or in other words

$$L(q) \geq \psi(q) + C_\psi, \tag{35}$$

where

$$C_\psi = \inf_\Omega(\log D - \psi) - c.$$

(Of course, at this stage, C_ψ could be $-\infty$, in which case Equation 35 is vacuous.)

By taking the function ψ_ϵ, say, we immediately get a lower bound $L \geq$ const. over Ω. (This is all we need for our main application below.) To prove Theorem 5 we need to make a more careful choice of the comparison function ψ. In fact the optimal choice of this comparison function leads to an interesting Monge-Ampère differential inequality which we will digress to explain. We consider, for a fixed point q the set of functions ψ satisfying conditions (1) and (2) and with $C_\psi > -\infty$. The optimal bound we can get from the argument above is given by the supremum over this set of functions ψ of $\psi(q) + C_\psi$. Changing ψ by the addition of a constant, we may suppose that $C_\psi = 0$. In other words we have $L(q) \geq \lambda(q) - c$, where

$$\lambda(q) = \sup_{\psi\in X}\psi(q)$$

and X is the set of functions which satisfy (1),(2) and in addition

$$\log D \leq \psi. \tag{36}$$

This becomes more familiar if we write $R = -\exp(-\psi)$. Then

$$R_{ij} = e^{-\psi}(\psi_{ij} - \psi_i\psi_j) = e^{-\psi} M_{ij}.$$

So condition (1) is simply requiring that R be a convex function. We have $D = e^{n\psi}\det(R_{ij})$ so Equation 36 becomes

$$\det(R_{ij}) \geq (-R)^{n-1}.$$

We can sum up the discussion in the following way. Given a domain Ω, let $\mathcal{R}_\Omega$ denote the set of negative convex functions R on Ω with

$$\lim_{x\to\partial\Omega} \frac{R(x)}{d(x)} = -\infty$$

and with

$$\det(R_{ij}) \geq (-R)^{n-1}.$$

Define a function ρ_Ω by

$$\rho_\Omega(x) = -\sup_{R\in\mathcal{R}_\Omega} R(x).$$

Then we have

Proposition 3 *In either Case 1 or Case 2, a solution $u \in \mathcal{S}_{\Omega,\sigma}$ of Abreu's equation with boundary data satisfies*

$$\det u_{ij} \geq \rho_\Omega^{-1} \left(\frac{\sup_\Omega A}{n}\right)^n.$$

in Ω.

This follows from the argument above and the asymptotic behaviour 32, 33 (In Case 2 we could strengthen the statement a little, taking account of the $\log d$ term in 33.)

We return from this digression to complete the proof of Theorem 5. For this we take co-ordinates $(x_1, \ldots, x_{n-1}, y)$ on $\mathbf{R}^n$ and consider the function

$$r(\underline{x}, y) = y^\alpha(\frac{b}{2}\underline{x}^2 - 1),$$

on a cylinder $Z = \{(\underline{x}, y) : |\underline{x}| \leq 1, 0 < y < 1\}$. Here $\alpha \in (0,1)$ is given, as in the statement of the Theorem, and $b > 0$ will be specified later. We will choose

b with $b < 2$ so that r is negative on Z. Straightforward calculation shows that r is convex on Z provided $(1-\alpha)(1-\frac{b}{2}) \geq \alpha b$. This is equivalent to

$$\frac{2-b}{2+b} > \alpha,$$

so, whatever α is given, we can choose b so small that the condition holds. Then

$$\frac{\det(r_{ij})}{(-r)^{n-1}} = \alpha y^{\alpha-2}\left((1-\alpha) - \alpha\frac{\underline{x}^2}{1-\frac{b}{2}\underline{x}^2}\right)(\frac{b}{1-\frac{b}{2}\underline{x}^2})^{n-1},$$

which is bounded below on Z so $\det(r_{ij}) \geq C(-r)^{n-1}$ throughout Z for some fixed positive C. Writing $R = C^{-1}r$ we get a function with $\det R_{ij} \geq (-R)^{n-1}$ on Z and with $-R = O(y^\alpha)$ as $y \to 0$.

Now one checks first that the inequality to be proved is stable under rescaling of the domain Ω. Then given a point p in our domain Ω we can obviously suppose without loss of generality that Ω lies in the set Z (for some suitable K depending on Ω), that the origin lies in $\partial\Omega$ and that the minimum distance to the boundary of Ω from p is achieved at the origin. Then the function R constructed above (or, more precisely, its restriction to Ω) lies in the set $\mathcal{R}_\Omega$, since R/d tends to $-\infty$ on $\partial\Omega$, and the result follows from Proposition 5.

It should be possible to sharpen this bound in various ways, and this is related to the Monge-Ampère Dirichlet problem for convex, negative, functions R on Ω:

$$\det R_{ij} = (-R)^{n-1} \quad R|_{\partial\Omega} = 0.$$

4.2 An upper bound on the determinant

In this subsection we will modify the method of Trudinger and Wang in [12], [13] to obtain an upper bound on the determinant of (u_{ij}). The argument applies in any dimension but the result we obtain requires additional information on "modulus of convexity" of u for its application, see (5.1) below.

Consider a bounded domain with smooth boundary D in $\mathbf{R}^n$ and a smooth convex function u on D satisfying $S(u) = A$, with $u < 0$ in D and $u \to 0$ on ∂D. We suppose that u is smooth up to the boundary of D. We consider Euclidean metrics g_{ij} on $\mathbf{R}^n$ with determinant 1 and for each such metric let

$$C_g = \max_D g^{ij} u_i u_j.$$

Now let

$$C = \min_g C_g.$$

This defines an invariant C of the function u on D. Another way of expressing the definition, is that $\omega_n C^n$ is least volume of an ellipsoid containing the image of D under the map $\nabla u : D \to \mathbf{R}^n$, where ω_n is the volume of the unit ball in $\mathbf{R}^n$.

Theorem 6 *In this situation*

$$(\det(u_{ij}))^{1/n} \leq \left(\frac{5}{2} + \frac{aM}{2n}\right) eC\ (-u)^{-1},$$

in D*, where*

$$a = \max(0, -\min_D A) \quad , \quad M = \max_D(-u).$$

To prove this we let g be the metric with $C_g = C$ and consider the function

$$f = -L - n\log(-u) - \alpha g^{ij} u_i u_j$$

on U, where α is a positive constant to be fixed later. Thus f tends to $+\infty$ on ∂D and there is a point p in D where f attains its minimum. At the minumum $f_i = 0$ and $u^{ij} f_{ij} \geq 0$. The first of these gives

$$L_i = -\frac{n u_i}{u} - 2\alpha g^{pg} u_p u_{qi}. \tag{37}$$

The second gives

$$0 \leq -u^{ij} L_{ij} - \frac{n^2}{u} + \frac{n}{u^2} u^{ij} u_i u_j - 2\alpha(u^{ij} g^{pq} u_{pi} u_{qj} + u^{ij} g^{pq} u_p u_{qij}), \tag{38}$$

where we have used the fact that $u^{ij} u_{ij} = n$. The crucial step now is to observe that

$$u^{ij} g^{pq} u_{pi} u_{qj} = g^{pq} u_{pq},$$

the ordinary Euclidean Laplacian of u, in the metric g. We next use the form Equation 13 of Abreu's equation to see that

$$u^{ij} L_{ij} \geq u^{ij} L_i L_j + \underline{A}, \tag{39}$$

where $\underline{A} = \min_D S(u)$. Now Equation 37 gives, at the point p,

$$u^{ij} L_i L_j = u^{ij} \left(\frac{n^2}{u^2} u_i u_j + 4\alpha^2 g^{pg} g^{rs} u_p u_r u_{qi} u_{sj} + 4\frac{\alpha n}{u} g^{pq} u_i u_p u_{qj}\right),$$

which can be written in the simpler form

$$u^{ij} L_i L_j = \frac{n^2}{u^2} u^{ij} u_i u_j + 4\alpha^2 g^{pq} g^{rs} u_{pr} u_q u_s + 4\frac{\alpha n}{u} g^{pq} u_p u_q.$$

Our inequalities 38, 39 give

$$\begin{aligned} 0 \leq \quad & - \quad \underline{A} - \frac{n^2}{u^2}(u^{ij} u_i u_j) - 4\alpha^2 (g^{pq} u_{rs} u_p u_q) - \frac{4\alpha n}{u} g^{pq} u_p u_q - \frac{n^2}{u} + \\ & + \quad \frac{n}{u^2}(u^{ij} u_i u_j) - 2\alpha g^{pq} u_{pg} - 2\alpha u^{ij} g^{pq} u_{pij} u_q. \end{aligned} \tag{40}$$

Now recall that $L_q = u^{jk}u_{qjk}$, so Equation 37 gives

$$g^{pq}u^{ij}u_{pij}u_q = g^{pq}L_p u_q = -\frac{n}{u}g^{pq}u_p u_q - 2\alpha u_{pq}u_r u_s g^{pr}g^{qs}$$

This means that three of the terms in (40) cancel. There are also two terms involving the expression $u^{ij}u_i u_j$ which we can combine to get

$$0 \leq -\underline{A} - \frac{n^2}{u} + \frac{(n-n^2)}{u^2}(u^{ij}u_i u_j) - 2\frac{\alpha n}{u}g^{pq}u_p u_q - 2\alpha g^{pq}u_{pq}. \qquad (41)$$

Now use the fact that $n - n^2 \leq 0$ and

$$g^{pq}u_p u_q \leq C,$$

by the definition of C. Re-arranging, we obtain

$$\frac{g^{pq}u_{pq}}{n} \leq \left(\frac{n}{2\alpha} + C)\right)\left(\frac{-1}{u}\right) - \frac{\underline{A}}{2\alpha n},$$

where $\underline{A} = \min A$. By the definition of a and M we have

$$\frac{g^{pq}u_{pq}}{n} \leq \left(\frac{-1}{u}\right)\left(\frac{n}{2\alpha} + C + \frac{aM}{2\alpha}\right).$$

Now, since g^{pq} has determinant 1 we have (as in Equation 34)

$$e^L = \det(u_{ij}) \leq \left(\frac{g^{pq}u_{pq}}{n}\right)^n,$$

so, at the point p,

$$L \leq -n\log(-u) + n\log P,$$

where $P = \frac{n+aM}{2\alpha} + C$. Going back to the definition of f we obtain

$$f(p) \geq -\alpha g^{pq}u_p u_q - n\log P \geq -\alpha C - n\log P.$$

So at any other point of D, we also have$f \geq -\alpha C - n\log P$ which gives

$$\det(u_{ij})^{1/n} \leq \kappa C(-u)^{-1},$$

with $\kappa = \exp(\frac{\alpha C}{n} + \log P)$. Optimising the choice of α one finds that the least value of κ is

$$\kappa_{\min} = C\left(\sqrt{m + \frac{m^2}{4}} + \frac{m}{2} + 1\right)\exp\left(\sqrt{m + \frac{m^2}{4}} - \frac{m}{2}\right),$$

where $m = \frac{1}{2} + \frac{aM}{2n}$. Now $\kappa_{\min} \leq Ce(m+2)$ since $\sqrt{m + \frac{m^2}{4}} \leq \frac{m}{2} + 1$, and this gives the inequality stated in the Theorem (*i.e.* the statement is not quite the optimal result: notice that if $A \geq 0$, so $a = 0$, one gets $\kappa_{\min} = 2Ce^{1/2}$.)

5 Estimates for higher derivatives

5.1 The modulus of convexity and the proof of Theorem 1

Suppose we have upper and lower bounds on $\det(u_{ij})$ in the interior of Ω for a solution u of Abreu's equation, for any smooth function A and in any dimension. We can then obtain estimates on all derivatives, following the discussion of Trudinger and Wang [12], [13], in terms of a "modulus of convexity" of u. This uses the results of Caffarelli [3] and Caffarelli and Gutiérrez [4] and the method exploits the form Equation 14, $Q(F) = -A$, of Abreu's equation.

We introduce some notation which we will use more extensively in (5.4) below. For a smooth strictly convex function u on any open set $U \subset \mathbf{R}^n$ and a point $x \in U$ we let λ_x be the affine linear function defining the supporting hyperplane of u at x (*i.e.* $u - \lambda_x$ vanishes to first order at x). Then we define the function H_x on U by

$$H_x(y) = u(y) - \lambda_x(y).$$

(So H_x is the normalisation of u at x, in our previous terminology.) Thus $H_x(y) \geq 0$ with equality if and only if $x = y$. We can think of the function $H_x(y)$ of two variables x, y as a kind of "distance function" on U, although it need not satisfy the axioms of a metric. For a subset $S \subset U$ we put $H_x(S) = \inf_{y \in S} H_x(y)$.

Now return to the case where u is a convex function on Ω and $K \subset\subset K^+$ are compact convex subsets of Ω with $0 < \lambda \leq \det u_{ij} \leq \Lambda$ on K^+. Let

$$H(K, K^+) = \min_{x \in K} H_x(\partial K^+).$$

Caffarelli and Gutiérrez prove that there is an $\alpha \in (0, 1)$ such that for any solution f of the linear equation $Q(f) = -A$ and x, y in K

$$|f(y) - f(x)| \leq C H(K, K^+)^{-\alpha} |x - y|^\alpha$$

where C depends only on A, the supremum of $|\nabla u|$ over K^+ and the upper and lower bounds Λ, λ of $\det(u_{ij})$ over K. [The results of [4] are stated for the case when $A = 0$ but, according to Trudinger and Wang ([13], discussion following Lemma 2.4), the arguments go over to the inhomogeneous equation.] Thus in our situation, taking $f = F = \det(u_{ij})^{-1}$, if we have a positive lower bound on $H(K, K^+)$ we get a Holder estimate on F and hence on $\det(u_{ij})$. Then the results of Caffarelli in [3] give $C^{2,\alpha}$ bounds on u over K, depending again on the $H(K, K^+)$. This gives C^α control of the co-efficients of the linearised operator Q and we can apply the Schauder estimates ([7] Theorem 6.2) to get $C^{2,\alpha}$ control of F, and so on.

Further, if we have a lower bound on the "modulus of convexity" $H(K, K^+)$ we can apply Theorem 6, using the interior bound Corollary 3 on the first

derivatives of u, to obtain an upper bound on $\det(u_{ij})$ over K. Given a point x in K we set

$$\tilde{u} = H_x - \frac{1}{2}H(K, K^+)$$

and

$$D = \{y : \tilde{u}(y) \leq 0\},$$

so $D \subset K^+$ and we can apply Theorem 6 to $\tilde{u}$ to get an upper bound on $\det(u_{ij})$ at x. Combining this with the lower bound from (4.1) (and replacing K^+ by K) we can then feed into the the preceding argument. Thus the obstacle to proving a result like Theorem 1 in general dimensions n is the need to control the $H(K, K^+)$, for interior subsets $K \subset\subset K^+ \subset \Omega$. In dimension 2 we can again argue in parallel with Trudinger and Wang in [13]. The result of (4.1) gives a lower bound $\det(u_{ij}) \geq \lambda_{K^+} > 0$, over K^+, for a solution satisfying our boundary conditions. Then a result of Heinz [10], implies that in two dimensions,

$$H(K, K^+) \geq C\lambda_{K^+}^{1/2},$$

where the constant C depends on K^+, the distance from K to the boundary of K^+ and $\sup_{K^+} |\nabla u|$. Putting all these facts from the literature together, we arrive at a proof of our main Theorem 1.

As we explained in the Introduction, we will give below an alternative proof, in the case when A is constant and Ω is a polytope, which avoids the sophisticated analysis of the Caffarelli-Gutiérrez theory but employs some of their basic tools. The modulus of convexity will also enter in a crucial way in our argument. We close this subsection with two further remarks

- In the case when A is a constant and Ω is a polygon in $\mathbf{R}^2$ we can combine the lower bound Theorem 5 and the upper bound Theorem 4, to obtain

$$|\nabla u| \geq C d^{-\alpha}$$

  near the boundary of Ω, for some $C, \alpha > 0$. This gives easily a lower bound on $H(K, K^+)$ for suitable K, K^+ and means we can avoid appealing to the result of Heinz in this case.

- In two dimensions the main result of Heinz in [10] gives a $C^{1,\beta}$ bound on u, in terms of upper and lower bounds for $\det(u_{ij})$. In the case when A is a constant we can use Lemma 3, and the bound on the vector field v (Theorem 2) to obtain a $C^{,\beta}$ bound on $\det(u_{ij})$. This gives an alternative path, avoiding appeal to [4], but feeding into [3], in this case.

The strategy for our alternative proof is as follows. In (5.2) we introduce two tensors F, G depending on the 4th. order derivatives of the function u and related to the Riemann curvature and Ricci tensors of a certain Riemannian

metric. We show that our boundary conditions fix the natural L^2 norms of these tensors. Then we show in (5.2) that under suitable conditions the L^∞ norm of G (or F) controls the second derivatives of the function u, while in (5.3) we show that under suitable hypotheses, including control of the second derivatives of u, the L^∞ norm of G is controlled by the L^2 norm of F. Putting together these three ingredients we complete the proof in (5.4), using a scaling argument and some of the basic geometrical results of Caffarelli and Gutiérrez on the sections of a convex function.

5.2 Curvature identities and L^2 bounds

For a convex function u on an open set U in $\mathbf{R}^n$ we define a 4-index tensor by

$$F^{ab}_{kl} = -u^{ab}_{kl}.$$

We can raise and lower indices in the usual way, using the metric u_{ij}, setting

$$F^{abcd} = u^{ck}u^{dl}F^{ab}_{kl} \; , F_{ijkl} = u_{ia}u_{jb}F^{ab}_{kl}.$$

Lemma 4 *The tensor F_{ijkl} is symmetric in pairs of indices*

$$F_{ijkl} = F_{klij}.$$

In fact, calculation gives,

$$F_{ijkl} = u_{ijkl} - u^{\lambda q}(u_{kjq}u_{il\lambda} + u_{ik\lambda}u_{jlq})$$

which makes the symmetry apparent.

We now introduce a Riemannian metric, due to Guillemin [8], on the $2n$-dimensional manifold $U \times \mathbf{R}^n$, with co-ordinates η_i in the second factor;

$$g = u_{ij}dx^i dx^j + u^{ij}d\eta_i d\eta_j. \tag{42}$$

This is in fact a Kahler metric: complex co-ordinates and a Kahler potential are furnished by the Legendre transfrom construction. If $\xi_i = u_i$ are the usual transformed co-ordinates we set $z_i = \xi_i + \sqrt{-1}\;\eta_i$ so

$$dz_i = u_{ij}dx^j + \sqrt{-1}\; d\eta_i. \tag{43}$$

Lemma 5 *The curvature tensor of g is*

$$-F^{ijkl}dz_i d\overline{z}_k \otimes dz_j d\overline{z}_l.$$

Of course we can use (43) to express the curvature tensor entirely in terms of products of the dx^i and $d\eta_j$, avoiding the Legendre transform.

The proof of the Lemma is mainly a matter of notation. We use the standard fact that if the Kahler metric is expressed by a (Hermitian) matrix-valued

function H, in complex co-ordinates, then the curvature, viewed as a matrix of 2-forms, is $-\overline{\partial}(H(\partial(H^{-1})))$. In our case the matrix H has entries

$$H_{\lambda\mu} = \langle \frac{\partial}{\partial z_\lambda}, \frac{\partial}{\partial z_\mu} \rangle = u^{\lambda\mu}.$$

so $\Gamma = -H(\partial H^{-1})$ is the matrix of 1-forms with $(\lambda\mu)$ entry

$$-u^{\lambda\mu} \frac{\partial u_{\mu\nu}}{\partial \xi_\alpha} dz_\alpha.$$

Now use the fact that $\frac{\partial}{\partial \xi_\alpha} = u^{\alpha i} \frac{\partial}{\partial x^i}$ to write this as

$$\Gamma^\lambda_\mu = -u^{\lambda\nu} u^{\alpha i} u_{\mu\nu i} dz^\alpha,$$

and this is just

$$\Gamma^\lambda_\mu = u^{\lambda\alpha}_\mu dz_\alpha.$$

So the curvature, as a matrix of 2-forms, has λ, μ entry

$$\frac{\partial}{\partial \xi_\beta} \left(u^{\lambda\alpha}_\mu \right) d\overline{z}_\beta dz_\alpha.$$

Expressing the derivatives in terms of the x^i variables again, this just means that the $(1, 3)$ curvature tensor is

$$u^{\beta j} u^{\lambda\alpha}_{\mu j} dz_\lambda \otimes \frac{\partial}{\partial z_\mu} \otimes d\overline{z}_\beta dz_\alpha,$$

and lowering an index gives the stated formula for the $(0, 4)$ curvature tensor.

Of course we can also obtain Lemma 3 via Lemma 4 and the usual symmetries of the curvature tensor, but we preferred to give the direct calculation. In fact in what follows we will make little explicit use of the Riemannian metric g, and derive most of the formulae we need directly.

Now define a 2-tensor G by contracting F:

$$G^i_k = F^{ij}_{kj}, \tag{44}$$

and likewise G^{ik}, G_{ik}. If follows from Lemma 3 that the latter are symmetric tensors, and contracting in Lemma 4 shows that G is essentially equivalent to the Ricci tensor of the metric g. In terms of our vector field v,

$$G^i_k = v^i_k.$$

A further contraction yields the scalar invariant $S = G^i_i$ of u which is of course the term appearing in Abreu's equation and which corresponds to the scalar curvature of g.

It is well-known in Kahler geometry that on a compact Kahler manifold the L^2-norms of the scalar curvature, the Ricci curvature and the full curvature tensor give essentially equivalent functionals on the metrics in a given Kahler class: they are related by topological invariants of the data [5]. In our setting we do not necessarily have a compact Kahler manifold available so we will develop the corresponding theory directly.

On account of the symmetry of Lemma 3, the standard square-norm of the tensor F, using the metric u_{ij} is

$$|F|^2 = F^{ij}_{kl} F^{kl}_{ij},$$

and this is the same (up to a numerical factor) as the standard square-norm of the curvature tensor of g. Similarly

$$|G|^2 = G^i_j G^j_i,$$

is essentially the square-norm of the Ricci tensor. We consider a 1-parameter family of functions $u(t)$ with

$$\frac{d}{dt} u|_{t=0} = \epsilon,$$

so ϵ is a function on U. We write E^{ij} for the t-derivative of u^{ij} at $t=0$, so

$$E^{ij} = -u^{ia} \epsilon_{ab} u^{bj}.$$

Proposition 4 *The time derivatives at $t = 0$ satisfy:*

$$\frac{d}{dt}\left(|F|^2 - |G|^2\right) = 2Z^i_i$$

and

$$\frac{d}{dt}\left(|G|^2 - S^2\right) = 2W^i_i$$

where

$$Z^i = -E^{jl}_k F^{ik}_{jl} + E^{ij}_k G^k_j,$$

and

$$W^i = -E^{jk}_j G^i_k + S E^{ji}_j.$$

The proofs are straightforward calculations.

Now return to the setting of our convex domain Ω in $\mathbf{R}^n$, where we suppose we are in Case 1 with Ω a polytope. Any pair u_0, u_1 of functions in $\mathcal{S}_\Omega$ can be joined by a smooth path u_t (for example a linear path) in $\mathcal{S}_\Omega$ For $\delta > 0$ let Ω_δ be an interior domain with boundary a distance δ from the boundary of Ω. Then the time derivative of

$$\frac{1}{2} \int_{\Omega_\delta} |F|^2 - S^2$$

is the boundary integral

$$\int_{\partial\Omega_\delta} Z^i + W^i. \tag{45}$$

Lemma 6 *The boundary integral (45) tends to* 0 *with* δ*, uniformly over the parameter* $t \in [0,1]$

Given this we obtain

Corollary 5 *In the case when* Ω *is a polytope there is an invariant* $\chi(\Omega,\sigma)$ *such that for any* $u \in \mathcal{S}_{\Omega,\sigma}$

$$\int_\Omega |F|^2 - S^2 = \chi(\Omega,\sigma).$$

It is easy to see that $\chi(\Omega,\sigma)$ is a tame function of Ω,σ. Now if $u \in \mathcal{S}_{\Omega,\sigma}$ has $S(u)$ constant then the value of the constant is fixed by Equation **??** and we see that

$$\int_\Omega |F|^2 = \chi(\Omega,\sigma) + \frac{\mathrm{Vol}(\partial\Omega,\sigma)^2}{\mathrm{Vol}(\Omega)},$$

is a tame function of Ω,σ.

To prove Lemma 5 we go back to Proposition 2. Note first that, in an adapted co-ordinate system, the matrix u^{ij} is smooth up to the boundary so the same is true for F^{ij}_{kl} and for the time derivative E^{ij}. (Actually, in Proposition 2 we chose a special adapted co-ordinate sytem to diagonalise the Hessian of the function in the variables x_i for $i > p$ but it is easy to see that the same conclusions hold without this restriction.) Thus the vector field Z, W are smooth up to the boundary and we simply need to show that the normal components vanish on the boundary. Thus we can restrict to a neighbourhood of a point in an $(n-1)$-dimensional face of the boundary, i.e. with $p = 1$. Now, according to Proposition 2, the entries u^{1j} are all products of x^1 with smooth functions so $u^{1j}_k = 0$ for all j and $k > 1$. Thus

$$E^{1j}_k = 0 \quad k > 1. \tag{46}$$

The x^1 derivative of u^{11} is fixed by the boundary measure so

$$E^{11}_1 = 0. \tag{47}$$

As for the second derivatives we have

$$F^{1j}_{kl} = -u^{1j}_{kl} = 0 \ \text{ for } k,l > 1 \tag{48}$$

and

$$F^{11}_{j1} = -u^{11}_{j1} = 0 \tag{49}$$

since the boundary measure is constant. It is now completely straightforward to check that (46),(47),(48),(49) imply that all the terms vanish in the sums

$$Z^1 = -E^{jl}_k F^{1k}_{jl} + E^{1j}_k G^k_j \ \ , \ \ W^1 = -E^{jk}_j G^1_k + S E^{j1}_j.$$

5.3 Estimates within a section

We begin by considering a convex function u on a bounded open set, with smooth boundary, $D \subset \mathbf{R}^n$. We assume that u is smooth up to the boundary, that $u<0$ in D and that u vanishes on ∂D. Suppose that, at each point of D, we have

$$u \geq -c_0, |G| \leq c_1, |v|_{\mathrm{Euc}} \leq c_2, |\nabla u|_{\mathrm{Euc}} \leq c_3,$$

for some $c_0, c_1, c_2, c_3 > 0$. Here G denotes the tensor G_{ij} introduced above and $|G|$ is the natural norm computed in the metric defined by u, that is

$$|G|^2 = G^i_j G^j_i = G_{ij} G_{ab} u^{ia} u^{jb}.$$

On the other hand the quantities $|v|_{\mathrm{Euc}}, |\nabla u|_{\mathrm{Euc}}$ are the norms of the vector field v^i and the derivative u_i computed with respect to the standard Euclidean metric on $\mathbf{R}^n$.

Proposition 5 *There is a constant K, depending only on n, c_0, c_1, c_2, c_3 such that $(u_{ij}) \leq K|u|^{-1}$ on D.*

The proof is a straightforward variant of Pogorelov's estimate for solutions of the Monge-Ampère equation $\det(u_{ij}) = 1$, see [9] Chapter 4. (It is also similar to the proof in (4.2) above.) It obviously suffices to estimate the second derivative u_{11} along the co-ordinate axis, and we set

$$f = \log(u_{11}) + \frac{u_1^2}{2}.$$

Now consider the function $\log(-u) + f$ which tends to $-\infty$ on the boundary so has an interior maximum. At this maximum point we have

$$\frac{u_i}{u} + f_i = 0 \tag{50}$$

and

$$P(-\log(u) + f) \leq 0$$

where $P(\phi) = (u^{ij}\phi_i)_j$. Now

$$P\log(u_{11}) = \left(\frac{u^{ij}u_{11i}}{u_{11}}\right)_j$$

which gives

$$P\log(u_{11}) = \frac{u^{ij}u_{11ij}}{u_{11}} + \frac{u^{ij}_j u_{11i}}{u_{11}} - \frac{u^{ij}u_{11i}u_{11j}}{u_{11}^2}. \tag{51}$$

Now

$$u^{ij}u_{11ij} = \left(u^{ij}u_{11i}\right)_j - u^{ij}_j u_{11i} = -\left(u^{ij}_1 u_{i1}\right)_j - u^{ij}_j u_{11i}.$$

Expanding out the derivatives we get

$$u^{ij}u_{11ij} = -u^{ij}_{1j}u_{i1} + u^{ia}u^{jb}u_{1ab}u_{1ij} - u^{ij}_{j}u_{11i}. \tag{52}$$

Here we recognise the expression $-u^{ij}_{1j}u_{i1}$ as the co-efficient G_{11} of the tensor G. (In fact the manipulation above is essentially the familiar identity, in Riemannian geometry, for $\Delta|d\phi|^2$ involving the Ricci tensor, where ϕ is a harmonic function, but we have preferred to do the calculation directly.) Substituting (52) into (51), two terms cancel and we get

$$P\log(u_{11}) = \frac{G_{11}}{u_{11}} + \frac{1}{u_{11}}u^{ia}u^{jb}u_{1ab}u_{1ij} - \frac{1}{u_{11}^2}u^{ij}u_{11i}u_{11j}. \tag{53}$$

Now simple calculations give

$$P(\frac{u_1^2}{2}) = u_{11}, \tag{54}$$

$$P(\log(-u)) = \frac{n - v^i u_i}{u} - \frac{u^{ij}u_i u_j}{u^2}. \tag{55}$$

So we conclude that, at the maximum point,

$$\frac{n - v^i u_i}{u} + u_{11} + \frac{G_{11}}{u_{11}} + (A) - (B) - (C) \leq 0,$$

where (A), (B) and (C) are the positive quantities

$$(A) = \frac{1}{u_{11}}u^{ia}u^{jb}u_{1ab}u_{1ij}, \tag{56}$$

$$(B) = \frac{1}{u_{11}^2}u^{ij}u_{11i}u_{iij}, \tag{57}$$

$$(C) = \frac{1}{u^2}u^{ij}u_i u_j. \tag{58}$$

We now use the condition (50) on the first derivatives. We may suppose that u_{ij} is diagonal at the given maximum point, with diagonal entries $u_{ii} = \lambda_i$ say. Then we have for $i \neq 1$

$$\frac{u_i}{u} = -\frac{u_{11i}}{u_{11}}.$$

The term (C) is

$$(C) = \frac{1}{u^2}\sum \lambda_i^{-1}u_i^2.$$

We write this as

$$(C) = \frac{u_1^2}{u^2 u_{11}^2} + (C)',$$

where

$$(C)' = \frac{1}{u^2}\sum_{i>1}\lambda_i^{-1}u_i^2 = \sum_{i>1}\frac{1}{\lambda_i\lambda_1^2}u_{11i}^2.$$

On the other hand, computing at this point, the other sums become

$$(A) = \sum_{ij}\frac{1}{\lambda_1\lambda_i\lambda_j}u_{1ij}^2, \tag{59}$$

$$(B) = \sum_{i}\frac{1}{\lambda_i\lambda_1^2}u_{11i}^2. \tag{60}$$

Once sees from this that

$$(A) - (B) - (C)' \geq 0$$

so we conclude that at the maximum point

$$\frac{n - u_i v^i}{u} + u_{11} + \frac{G_{11}}{u_{11}} - \frac{u_1^2}{u^2 u_{11}} \leq 0. \tag{61}$$

This should be compared with the more standard calculation, for solutions of the Monge-Ampère equation $\det(u_{ij}) = 1$, when v and G vanish, so two terms (61) are absent.

Now, computing the norm in the diagonal basis,

$$|G|^2 = \sum \lambda_i^2 G_{ii}^2 \geq \left(\frac{C_{11}}{u_{11}}\right)^2,$$

so

$$|\frac{G_{11}}{u_{11}}| \leq c_1.$$

Clearly $n - u_i v^i \leq n + c_2 c_3$. Then (61) gives

$$\frac{N}{u} + u_{11} - \frac{u_1^2}{u^2 u_{11}} \leq 0,$$

where $N = n + c_2 c_3 + c_1 c_0$. Thus we can adapt the argument from the standard case, replacing n by N. If h is the function

$$h = \exp(\log(-u) + f) = |u| u_{11} \exp(\frac{u_1^2}{2}),$$

we see that at the maximum point for h

$$h^2 - N\exp(\frac{u_1^2}{2})h - u_1^2\exp(u_1^2) \leq 0,$$

which gives

$$h_{max} \leq \exp(\frac{u_1^2}{2})\frac{1}{2}\left(N + \sqrt{N^2 + 4u_1^2}\right).$$

This gives

$$u_{11} \leq \frac{K}{|u|},$$

where

$$K = \frac{1}{2}\exp(\frac{c_3^2}{2})\left(N + \sqrt{N^2 + 4c_3^2}\right).$$

5.4 Yang-Mills estimate

In this subsection we consider a solution of the equation $S(u) = A$ with A a *constant*. Roughly speaking, we show that when the dimension n is 2 and given uniform bounds on the Hessian of u, the L^2 norm of the tensor F (the "Yang-Mills functional") controls the L^∞ norm.

Proposition 6 *Suppose u is a convex function on the unit disc D in $\mathbf{R}^2$, smooth up to the boundary, which satisfies the equation $S(u) = A$ where A is constant. Set*

$$\mathcal{E} = \int_D |F|^2.$$

If the Hessian of u is bounded above and below

$$K^{-1} \leq (u_{ij}) \leq K$$

then

$$|G(0)|^2 \leq \kappa(\mathcal{E} + \mathcal{E}^3)$$

and for any $p > 1$

$$\int_{\frac{1}{2}D} |F|^p \leq \kappa_p(\mathcal{E} + \mathcal{E}^3)^p$$

where κ depends only on K and $\sup_D |v|$ and κ_p depends on K and p.

The proof makes use of the Sobolev inequalities. In dimension 2 we have

$$\|\phi\|_{L^p} \leq C_p \|\nabla\phi\|_{L^2}$$

for any p and compactly supported ϕ on D (here all norms are the standard Euclidean ones). The proof can be modified to give a similar result in dimension 3 and extended to give information when $n = 4$ provided $\mathcal{E}$ is sufficiently small, in the manner of Uhlenbeck[14] and, still more, Anderson [2] and Tian and Viaclovsky[11]. In the proof we make more use of the Riemannian metric g on $D \times \mathbf{R}^2$ defined by the convex function u.

The idea of the proof is to exploit the fact that the tensors F and G satisfy quasi-linear elliptic equations. To derive these we use the interpretation of these tensors as (essentially) the Riemannian curvature and Ricci tensor of the metric g see also [11]. (Of course there is no particular difficulty in deriving these equations directly, without explicit reference to Riemannian and Kahler geometry, but the derivation involves manipulating sixth order derivatives of the function u.)

Lemma 7 *Suppose that g is a Kahler metric of constant scalar curvature. Then the Riemann tensor* Riem *and Ricci tensor* Ric *of g satisfy*

$$\nabla^*\nabla \mathrm{Ric} = \mathrm{Riem} * \mathrm{Ric},$$

$$\nabla^*\nabla \mathrm{Riem} = \mathrm{Riem} * \mathrm{Riem} + 2\nabla'\nabla''\mathrm{Ric},$$

where $*$ denotes appropriate natural algebraic bilinear forms and ∇', ∇'' are the $(1,0)$ *and* $(0,1)$ *components of the covariant derivative.*

To derive these identities we can consider more generally the curvature tensor Φ of a holomorphic vector bundle E over a Kahler manifold M. Then we have $\overline{\partial}$-operators

$$\overline{\partial} : \Omega^{p,q}(\mathrm{End}E) \to \Omega^{p,q+1}(\mathrm{End}E).$$

The Laplacians $\nabla^*\nabla$ and $\Delta_{\overline{\partial}} = 2(\overline{\partial}^*\overline{\partial} + \overline{\partial}\overline{\partial}^*)$ differ by a Weitzenbock formula involving the curvature of the bundle and base manifold. Since $\overline{\partial}\Phi = 0$ we have

$$\Delta_{\overline{\partial}}\Phi = 2\overline{\partial}\overline{\partial}^*\Phi.$$

Now the Kahler identities give

$$\overline{\partial}^*\Phi == i\nabla'(\Lambda\Phi),$$

where $\Lambda : \Omega^{1,1}(\mathrm{End}E) \to \Omega^0(\mathrm{End}E)$ is the trace on the form component. Thus $\nabla^*\nabla\Phi$ is equal to $2i\overline{\partial}\nabla(\Lambda\Phi) = 2i\nabla''\nabla'(\Lambda\Phi)$ plus a bilinear algebraic term involving Φ and the curvature of the base manifold.

We apply this first to the bundle $E = \Lambda^n TM$, the anticanonical line bundle. In this case Φ is essentially the Ricci tensor and $\Lambda\Phi$ is the scalar curvature. So the constant scalar curvature condition gives $\Delta_{\overline{\partial}}\Phi = 0$ and the first formula follows from the discussion above. The second formula does not use the constant scalar curvature condition. It follows from the discussion above taking $E = TM$, when Φ is the Riemann curvature tensor and $\Lambda\Phi$ is the Ricci tensor.

In our situation we deduce that F, G satisfy equations which we write, rather schematically as

$$\nabla^*\nabla G = F * G \tag{62}$$

$$\nabla^*\nabla F = F * F + 2\nabla'\nabla'' G \tag{63}$$

The strategy now is to take the L^2 inner product of these identities with suitable compactly-supported tensor fields. It is convenient here to regard g as a metric on the 4-manifold $D\times T^2$ where T^2 is the torus $\mathbf{R}^2/\mathbf{Z}^2$. However all our data will depend only on the D variables and the torus factor plays an entirely passive role: the push-forward of the volume form on the 4-manifold to D is the standard Lebesgue meausure and the Laplacian of g, applied to T-invariant functions is the operator P. Notice the crucial fact that for fixed K the metric g is uniformly equivalent to the standard Euclidean metric.

To prove the first part of Proposition 6 we fix a standard cut-off function β on the disc, equal to 1 in a neighbourhood of 0, and take the L^2 inner product with $\beta^2 G$ on either side of (62). This yields

$$\int_{D\times T^2} \nabla(\beta^2 G).\nabla G \leq C\int_D \beta^2 |G|^2 |F|.$$

(In this proof C will denote an unspecified constant, changing from line to line.) We have

$$\nabla(\beta^2 G).\nabla G = |\nabla(\beta G)|^2 - |G|^2|\nabla\beta|^2$$

so we get

$$\int_{D\times T^2} |\nabla(\beta G)|^2 \leq C\int_D |G|^2 \sqrt{\nabla}\beta|^2 + \beta^2|G|^2|F|. \tag{64}$$

Applying the Sobolev inequality to $|\beta G|$ and Cauchy-Schwartz we obtain

$$\|\beta G\|_{L^6}^2 \leq C\left(\|G\|_{L^2}^2 + \|\beta G\|_{L^4}^2 \|f\|_{L^2}\right).$$

(Here we are using the fact that, for fixed K, the metric g is uniformly equivalent to the Euclidean metric, so we can transfer the standard Sobolev inequalities to our setting.) Using

$$\|\beta G\|_{L^4}^4 \leq \|\beta G\|_{L^2}\|\beta G\|_{L^6}^3,$$

we get

$$\|\beta G\|_{L^6}^2 \leq C\left(\|G\|_{L^2}^2 + \|\beta G\|_{L^6}^{3/2}\|F\|_{L^2}\|\beta G\|_{L^2}^{1/2}\right) \leq C\left(\mathcal{E}^2 + \mathcal{E}^{3/2}\|\beta G\|_{L^6}^{3/2}\right).$$

This yields

$$\|\beta G\|_{L^6} \leq C(\mathcal{E}+\mathcal{E}^3).$$

Thus we have an L^6 bound on G in a neighbourhood of 0 and this gives an $L^{3/2}$ bound on $F*G$ over this neighbourhood. Going back to (64), we also have an L^2 bound on the derivative of βG:

$$\|\nabla(\beta G)\|_{L^2} \leq C(\mathcal{E}+\mathcal{E}^3).$$

We introduce another cut-off function γ, supported on the neighbourhood where $\beta = 1$. Then we have

$$\Delta(\gamma|G|) \geq |\nabla^*\nabla(\gamma G)| \geq |\Delta\gamma||G| + |\nabla\gamma||\nabla G| + C\gamma|F||G|.$$

Now

$$\Delta\gamma = u^{ij}\gamma_{ij} + v^i\gamma_i.$$

So $|\Delta\gamma|$ satisfies a bound, depending on $\sup|v|$ and we have

$$\Delta(\gamma|G|) \geq \sigma,$$

say where

$$\|\sigma\|_{L^{3/2}} \leq C(\mathcal{E} + \mathcal{E}^3).$$

Then we can apply Theorem 8.15 of [7] (proved by the Moser iteration technique) to obtain the desired bound on $|G|$ at 0.

For the second part of Proposition 6 we operate with the equation (62). We take the L^2 inner product with $\gamma^2 F$ and integrate by parts. For the term involving the second derivatives of G we write

$$\int_{D\times T^2} \gamma^2 F.\nabla''\nabla' G = \int_{D\times T^2} (\nabla'')^*(\gamma^2 F).\nabla' G.$$

This yields

$$\int_{D\times T^2} |\nabla(\gamma F)|^2 \leq C\int_{D\times T^2} \gamma^2|F|^3 + |F|^2|\nabla\gamma|^2 + \gamma|\nabla(\gamma F)||\nabla G + \gamma|\nabla\gamma||F||\nabla G|.$$

Then, using the Sobolev inequality as before and re-arranging we get an L^2 bound on the derivative of γF near 0 which, in dimension 2, gives the required L^p bound (since we can suppose that $\gamma = 1$ on the disc $\frac{1}{2}D$).

5.5 Rescaling sections

In this subsection we will bring together the three ingredients established above to obtain a pointwise bound on the tensor G over compact subsets of Ω. This involves rescaling the geometric data. A rescaling argument of this kind, using balls determined by the Riemannian distance function, would be fairly standard. However there are difficulties in carrying this through unless one can establish some control of the injectivity radius, or something similar. We get around this difficulty by using Caffarelli's theory of the "sections" of a convex function, these taking the place of geodesic balls.

Recall from (5.1) that if u is a smooth, strictly convex function on an open set $U \subset \mathbf{R}^n$ and x, y is a point in U we have defined $H_x(y) \geq 0$, vanishing if and only if $x = y$. For $t \geq 0$ the *section* $S_x(t)$ at x and level t is the set

$$S_x(t) = \{y \in U : H_x(y) \leq t\}.$$

We will use three results about these sections, or equivalently the functions H_x, taken from [9]. For each of these results we suppose that the determinant of the Hessian satisfies upper and lower bounds

$$0 < \lambda \leq \det(u_{ij}) \leq \Lambda,$$

throughout U, and the constants c_i below depend only on λ, Λ. Recall that a convex set K in $\mathbf{R}^n$ is *normalised* if

$$\alpha_n B_n \subset K \subset B_n \tag{65}$$

where $\alpha_n = n^{-3/2}$. Any compact convex set can be mapped to a normalised set by an affine-linear transformation ([9], Theorem 1.8.2).

Proposition 7 *1.* ([9] page 50, Corollary 3.2.4) *Suppose that $S_x(t)$ is compact. Then*

$$c_1 t^{n/2} \leq \text{Vol } (S_x(t)) \leq c_2 t^{n/2}. \tag{66}$$

2. ([9] page 55, Corollary 3.3.6 (i)) *Suppose that $S_x(t)$ is compact and normalised. Then for $y \in S_x(t/2)$*

$$d(y, \partial S_x(t)) \geq c_3 > 0. \tag{67}$$

3. ([9] page 55, Theorem 3.3.7) *Suppose that $S_x(2t)$ is compact. Then if $H_x(y) \leq t$ and $H_x(z) \leq t$ we have $H_y(z) \leq c_4 t$.*

4. ([9] page 57, Theorem 3.3.8) *Suppose that $S_x(t)$ is compact and normalised. Then $S_x(t/2)$ contains the Euclidean ball of radius c_5 centred on x.*

The third result can be seen as a substitute for the triangle inequality, if one views $H_x(y)$ as a defining a notion of "distance" in U. (The assumption that $S_x(2t)$ be compact does not appear explicitly in [9], where it is assumed that $U = \mathbf{R}^n$ and all sections are compact, but a review of the proof shows that this is the hypothesis needed for our situation.)

We will now discuss the scaling behaviour of the tensors we have associated to a convex function u. For $t > 0$ and $T = T^a_j \in SL(n, \mathbf{R})$ we set

$$\tilde{u}(x) = t^{-1} u(\sqrt{t} x) \ , \ u^*(x) = \tilde{u}(Tx) = t^{-1} u(\sqrt{t} T^a_j x^j)$$

We write $S = S^j_a$ for the inverse matrix of $T = T^a_j$. Then we have

$$\begin{aligned} \det(u_{ij}) &= \det(\tilde{u}_{ij}) = \det(u^*_{ij}); && (68)\\ |F^*| &= |\tilde{F}| = t|F| && (69)\\ |G^*| &= |\tilde{G}| = t|G| && (70)\\ (v^*)^j &= S^j_a \tilde{v}^a = \sqrt{t} S^j_a v^a. && (71) \end{aligned}$$

Here we have an obvious notation, in which for example $\tilde{v}$ and v^* refer to the vector fields obtained from the convex functions $\tilde{u}, u^*$. The verification of all these identities is completely elementary.

With these prelminaries in place we can proceed to our main argument. From now on we fix $n = 2$ and suppose that u is a convex function on $\Omega \subset \mathbf{R}^2$ which satisfies Abreu's equation $S(u) = A$ with constant A. We set

$$E = \int_\Omega |F|^2$$

$$\rho = \sup_\Omega |v|_{Euc}.$$

Let $K \subset\subset K^+$ be compact subsets of Ω and suppose that

$$0 < \lambda \leq \det(u_{ij}) \leq \Lambda,$$

on K^+. For x in K we recall that

$$H_x(\partial K^+) = \min_{y \in \partial K^+} H_x(y),$$

and we let

$$\delta = H(K, K^+) = \min_{x \in K} H_x(\partial K^+).$$

We also put

$$D = \max_{x \in K} H_x(\partial K).$$

Finally, we define a function Φ on K by

$$\Phi(x) = |G(x)| H_x(\partial K),$$

and let

$$M = \max_{x \in K} \Phi(x).$$

Theorem 7 *There is a constant μ depending only on $\lambda, \Lambda, E, \rho, \delta, D$ such that $M \leq \mu$.*

To prove this we may obviously suppose that $M > 2$. We consider a point x_0 where Φ attains its maximum value M and set $t = |G(x_0)|^{-1}$. Then $M > 2$ implies that $2t < H_{x_0}(\partial K)$ so the section $\Sigma = S_{x_0}(t)$ lies in K. Hence Σ is compact in Ω. Our first goal is to show that $|G|$ over Σ is controlled by t^{-1}, i.e. by its value at x_0

Now we may suppose that

$$M \geq \max(2c_4^2, 4c_4 D\delta^{-1}),$$

where $c_4 > 1$ is the constant of (3) in Proposition 7, depending on the given bounds λ, Λ. This means that if we define ϵ to be

$$\epsilon = \min(1/2, \frac{c_4 \delta}{2D})$$

we have $M \geq c_4^2/\epsilon$.

We will now make two applications of the inequality in (3) of Proposition 7. For the first we simply observe that in fact $S_{x_0}(2t)$ lies in K and *a fortiori* in K^+. Hence this section is compact. Then for $y \in S_{x_0}(t)$ we have $H_{x_0}(y) \leq t$ and trivially $H_{x_0}(x_0) = 0 \leq t$, so we deduce that $H_y(x_0) \leq c_4 t$.

Now set $r = H_{x_0}(\partial K)$ so $t = r/M$. We claim that if y is in $S_{x_0}(t)$ we have

$$H_y(\partial K) \geq \frac{r\epsilon}{c_4}. \tag{72}$$

To see this, the result above yields

$$H_y(x_0) \leq c_4 t = c_4 r/M,$$

and $c_4 r/M \leq r\epsilon/c_4$ since we have arranged that $M \geq c_4^2/\epsilon$. Set $\tau = r\epsilon/c_4$. For $y \in S_{x_0}(t) \subset K$ we have

$$H_y(\partial K^+) \geq \delta \geq \frac{2\epsilon D}{c_4} \geq \frac{2\epsilon r}{c_4} = 2\tau. \tag{73}$$

Suppose that (72) is not true, so there is a point $z \in \partial K$ with $H_y(z) \leq \tau$. Then, by (73), $S_y(2\tau)$ lies in K^+ and $H_y(x_0) \leq \tau$ so (3) in Proposition 7 would give

$$H_{x_0}(z) \leq c_4 \tau = r\epsilon \leq r/2,$$

a contradiction to $H_{x_0}(\partial K) = r$.

Now from $H_y(\partial K) \leq r\epsilon/c_4$ and the definition of M we obtain our first goal:

$$|G(y)| \leq \frac{c_4}{\alpha}|G(x_0)| \tag{74}$$

for all $y \in \Sigma = S_{x_0}(t)$.

We now invoke our scaling construction for the restriction of u to Σ. We fix the real parameter t to be as above. We know that there is some $k > 0$ and an unimodular affine transformation T so that $kt^{-1/2}T^{-1}(\Sigma)$ is normalised. But we know from (1) of Proposition 7 that the volume of Σ lies between $c_1 t$ and $c_2 t$, so the volume of $kt^{-1/2}T^{-1}(\Sigma)$ lies between $k^2 c_1$ and $k^2 c_2$ hence

$$\frac{\pi}{8c_2} \leq k^2 \leq \frac{\pi}{c_1}. \tag{75}$$

Thus the convex set $\Sigma^* = t^{-1/2}T^{-1}(\Sigma)$ differs from its normalisation by a scale factor which is bounded above and below, so we can apply the results of (2) and (4) in Proposition 7, with a change in the constants depending on the above bounds for k (alternatively, we could normalise Σ^* by changing the definition of t).

Lemma 8 *There is a constant C depending only on λ, Λ such that*

$$|v^*|_{Euc} \leq C|v|_{Euc} \mathrm{Diam}(\Sigma).$$

To see this we may suppose that T is diagonal, with $T_1^1 = \lambda, T_2^2 = \lambda^{-1}, T_1^2 = T_2^1 = 0$. Then, by (71), v^* has components

$$(v^*)^1 = \sqrt{t}\lambda^{-1} v^1 \ , \ (v^*)^2 = \sqrt{t}\lambda v^2.$$

But Σ^* contains the disc of radius of radius $R = k^{-1}2^{-3/2}$ so $\Sigma = \sqrt{t}T(\Sigma^*)$ contains an ellipse with semi-axes $\sqrt{t}\lambda R, \sqrt{t}R\lambda^{-1}$. In particular this ellipse is contained in Ω so

$$\sqrt{t}\lambda, \sqrt{t}\lambda^{-1} \leq \frac{\mathrm{Diam}(\Omega)}{2R}.$$

Thus

$$|v^*|_{\mathrm{Euc}} \leq \frac{\mathrm{Diam}(\Omega)}{2R} |v|_{\mathrm{Euc}}. \tag{76}$$

The scaling behaviour (70) of the tensor G and the result (72) established above means that $|G^*| \leq c_4/\alpha$ over Σ^*. We now apply Proposition 5 to an interior set $\Sigma_0^* \subset \Sigma^*$. There is no loss of generality in supposing that $x_0 = 0$ and that u vanishes to first order at this point, so $\Sigma^* = \{y : u^*(y) \leq 1\}$. We define

$$\Sigma_0^* = \{y : u^*(y) \leq 3/4\},$$

and we set $u_0^* = u^* - 3/4$. The lower bound on the distance to $\partial\Sigma^*$ furnished by (2) of Corollary 7 gives, in an elementary way, a bound on the derivative of u_0^* over Σ_0^* and we can apply Proposition 5 to u_0^*, taking $D = \Sigma_0^*$. Since we have bounds on $|G^*|$ and $|v^*|_{\mathrm{Euc}}$ (the latter by Lemma 8), we get an upper bound on the Hessian of u_0^* over the further interior set

$$\Sigma_{-1}^* = \{y : u^*(y) \leq 1/2\}.$$

The bound on $\det(u_{ij}^*)$ then yields a lower bound on the Hessian over Σ_{-1}^*. We then use (4) of Proposition 7 to see that Σ_{-1}^* contains a disc Δ about 0 of fixed radius, and we can apply Proposition 6 to conclude that

$$|G^*(0)|^2 \leq C(\mathcal{E} + \mathcal{E}^3),$$

where

$$\mathcal{E} = \int_\Delta |F^*|^2 \leq C \int_{\Sigma^*} |F^*|^2.$$

(Here we use again the bound on $|v^*|_{\mathrm{Euc}}$.) Now the scaling behaviour of F implies that

$$\int_{\Sigma^*} |F^*|^2 = t \int_\Sigma |F|^2 \leq tE.$$

So we conclude that

$$|G^*(0)| \leq C(\sqrt{tE + t^3E^3},$$

for some constant C depending only on $\lambda, \Lambda, \delta, D, \rho$. But by construction $|G^*(0)| = 1$, so we obtain a lower bound on t, thus an upper bound on $|G(x_0)|$ and thence on M. This completes the proof of Theorem 7.

Given Theorem 7 it is straightforward to complete the proof of Theorem 1 in the case when A is a constant and $\Omega \subset \mathbf{R}^2$ is a polygon. By the discussion of uniform convexity in (5.1) above we can, for any compact set $K_0 \subset \Omega$, find further compact sets

$$K_0 \subset\subset K \subset\subset K^+$$

in Ω such that $\delta = H(K, K^+)$ and $\delta_0 = H(K_0, K)$ are bounded below by positive bounds depending continuously on the data. Then Theorem 7 implies that the tensor G is bounded on K. Covering K by a finite number of sections and applying the same argument as above we get upper and lower bounds on the Hessian (u_{ij}) over a neighbourhood of K. Further, we can apply the second part of Proposition 6 to get bounds on the L^p norm of the tensor F for any p. Since u_{ij} is bounded above and below these are equivalent to L^p bounds on the second derivatives of the matrix (u^{ij}). In dimension 2, L^p_2 functions are continuous, so we deduce L^p_2 bounds on the inverse matrix (u_{ij}). Thus we have uniform L^p_4—hence $C^{3,\alpha}$—bounds on u over a neighbourhood of K. From this point elementary methods suffice to bound all higher derivatives.

References

[1] M. Abreu *Kahler geometry of toric varieties and extremal metrics* Int. J. Math. **9** 641-651 1998

[2] M. T. Anderson *Orbifold compactness for spaces of Riemannian metrics and applications* Math. Annalen **331** 739-778 2005

[3] L.A. Caffarelli *Interior W^2_p estimates for solutions to the Monge-Ampére equation* Annals of Math. **131** 135-150 1990

[4] L.A. Caffarelli and C.E. Gutiérrez *Properties of the solutions of the linearised Monge-Ampère equation* Amer. Jour. Math. **119** 423-465 1997

[5] E. Calabi *Extremal Kahler metrics* In: *Seminar in Differential Geometry* (Ed. S-T.Yau) Annals of Math. Studies Princeton UP 1982

[6] S.K. Donaldson *Scalar curvature and stability of toric varieties* Jour. Differential Geometry **62** 289-349 2002

[7] D. Gilbarg and N.S. Trudinger *Elliptic Partial Differential Equations of Second Order* 2nd. Edition, Springer 1983

[8] V. Guillemin *Kaehler structures on toric varieties* Jour. Differential Geometry **40** 285-309 1994

[9] C.E. Gutiérrez *The Monge-Ampère equations* Birkhauser 2001

[10] E. Heinz *Über die Differentialungleichung* $0 < \alpha \leq tr - s^2 \leq \beta < \infty$ Math. Z. **72** 107-126 1959

[11] G. Tian and J. Viaclovsky *Bach-flat asymptotically locally Euclidean metrics* Preprint DG/0310302

[12] N.S. Trudinger and X. Wang *The Bernstein problem for affine maximal hypersurfaces* Inventiones Math. **140** 399-402 2000

[13] N.S Trudinger and X. Wang *Bernstein-Jörgens theorem for a fourth order partial differential equation* J. Partial Differential equations **15** 78-88 2002

[14] K. K. Uhlenbeck *Removeable singularities in Yang-Mills fields* Commun. Math. Phys. **83** 11-29 1982